

```

1 ATGGATTTTC TCCTGGCGCT GGTGCTGGTA TCCTCGCTCT ACCTGCAGGC
51 GGCCGCCGAG TTCGACGGGA GGTGGCCAG GCAAATAGTG TCATCGATTG
101 GCGTATGTCG TTATGGTGCG AGGATTGACT GCTGCTGGGG CTGGGCTCGC
151 CAGTCTTGGG GACAGTGTCA GCCTTTCTAC GTCTTAAGGC AGAGAATAGC
201 CAGGATAAGG TGCCAGCTCA AAGCTGTGTG CCAACCACGA TGCAAACATG
251 GTGAATGTAT CGGGCCAAAC AAGTGCAAGT GTCATCCTGG TTATGCTGGA
301 AAAACCTGTA ATCAAGACGA GCACATCCCA GCTCCTCTTG ACCAAGGCAG
351 TGAACAGCCT CTTTTCCAAC CCCTGGATCA CCAAGCCACA AGTTTGCCTT
401 CAAGGGATCT AAATGAGTGT GGCCTGAAGC CCCGGCCCTG TAAGCACAGG
451 TGCATGAACA CTTACGGCAG CTACAAGTGC TACTGTCTCA ACGGATATAT
501 GCTCATGCCG GATGGTTCCT GCTCAAGTGC CCTGACCTGC TCCATGGCAA
551 ACTGTCAGTA TGGCTGTGAT GTTGTTAAAG GACAAATACG GTGCCAGTGC
601 CCATCCCCTG GCCTGCAGCT GGCTCCTGAT GGGAGGACCT GTGTAGATGT
651 TGATGAATGT GCTACAGGAA GAGCCTCCTG CCCTAGATTT AGGCAATGTG
701 TCAACACTTT TGGGAGCTAC ATCTGCAAGT GTCATAAAGG CTTGATCTC
751 ATGTATATTG GAGGCAAATA TCAATGTCAT GACATAGACG AATGCTCACT
801 TGGTCAGTAT CAGTGCAGCA GCTTTGCTCG ATGTTATAAC GTACGTGGGT
851 CCTACAAGTG CAAATGTAAA GAAGGATACC AGGGTGATGG ACTGACTTGT
901 GTGTATATCC CAAAAGTTAT GATTGAACCT TCAGGTCCAA TTCATGTACC
951 AAAGGGAAAT GGTACCATTT TAAAGGGTGA CACAGGAAAT AATAATTGGA
1001 TTCCTGATGT TGAAGTACT TGGTGGCCTC CGAAGACACC ATATATTCCT
1051 CCTATCATT AACAAGGCC TACTTCTAAG CCAACAACAA GACCTACACC
1101 AAAGCCAACA CCAATTCCTA CTCCACCACC ACCACCACC CTGCCAACAG
1151 AGCTCAGAAC ACCTCTACCA CCTACAACCC CAGAAAGGCC AACCACCGGA
1201 CTGACAACTA TAGCACCAGC TGCCAGTACA CCTCCAGGAG GGATTACAGT
1251 TGACAACAGG GTACAGACAG ACCCTCAGAA ACCCAGAGGA GATGTGTTCA
1301 TTCCACGGCA ACCTTCAAAT GACTTGTTTG AAATATTTGA AATAGAAAGA
1351 GGAGTCAGTG CAGACGATGA AGCAAAGGAT GATCCAGGTG TTCTGGTACA
1401 CAGTTGTAAT TTTGACCATG GACTTTGTGG ATGGATCAGG GAGAAAGACA
1451 ATGACTTGCA CTGGGAACCA ATCAGGGACC CAGCAGGTGG ACAATATCTG
1501 ACAGTGTCGG CAGCCAAAGC CCCAGGGGGA AAAGCTGCAC GCTTGGTGCT
1551 ACCTCTCGGC CGCTCATGC ATTCAGGGGA CCTGTGCCTG TCATTCAGGC
1601 ACAAGGTGAC GGGGCTGCAC TCTGGCACAC TCCAGGTGTT TGTGAGAAAA
1651 CACGGTGCCC ACGGAGCAGC CCTGTGGGGA AGAAATGGTG GCCATGGCTG
1701 GAGGCAAACA CAGATCACCT TGCGAGGGGC TGACATCAAG AGCGTCGTCT
1751 TCAAAGGTGA AAAAAGGCGT GGTCACTG GGGAGATTGG ATTAGATGAT
1801 GTGAGCTTGA AAAAAGGCCA CTGCTCTGAA GAACGCTAA (SEQ ID NO:1)

```

FEATURES:

Start Codon: 1
Stop Codon: 1837

Homologous proteins:

Top 10 BLAST Hits:

| | Score | E |
|--|-------|-------|
| CRA 50000098943299 /altid=gi 15795193 /def=ref NP_277060.1 nep... | 1190 | 0.0 |
| CRA 46000103872918 /altid=gi 15128103 /def=gb AAK84391.1 AF3970... | 1147 | 0.0 |
| CRA 84000042916333 /altid=gi 15430248 /def=gb AAK96011.1 (AY03... | 1128 | 0.0 |
| CRA 84000042916331 /altid=gi 15430246 /def=gb AAK96010.1 (AY03... | 1084 | 0.0 |
| CRA 78000201656899 /altid=gi 14763663 /def=ref XP_040394.1 epi... | 397 | e-109 |
| CRA 163000000492958 /altid=gi 9506563 /def=ref NP_062270.1 EGF... | 396 | e-109 |
| CRA 157000140328010 /altid=gi 6752658 /def=gb AAF27812.1 AF1860... | 396 | e-109 |
| CRA 157000140333453 /altid=gi 13124888 /def=ref NP_056322.2 ep... | 395 | e-109 |
| CRA 1000682328934 /altid=gi 7512796 /def=pir T17324 hypothetical... | 385 | e-106 |
| CRA 60000046724650 /altid=gi 13929180 /def=ref NP_114014.1 fib... | 147 | 4e-34 |

FIGURE 1A

Blast hits to dbEST:

| CRA Number | gi Number | Score | Expect |
|---------------------|-------------|-----------------|--------|
| CRA 113000119195127 | gi 14391046 | 1068 bits (539) | 0.0 |
| CRA 110000024586745 | gi 12320908 | 866 bits (437) | 0.0 |
| CRA 87000000737360 | gi 7312275 | 842 bits (425) | 0.0 |
| CRA 113000119269070 | gi 14399233 | 831 bits (419) | 0.0 |
| CRA 113000119283425 | gi 14400823 | 827 bits (417) | 0.0 |
| CRA 113000119283416 | gi 14400822 | 785 bits (396) | 0.0 |
| CRA 112000057002060 | gi 14377756 | 680 bits (343) | 0.0 |
| CRA 147000029699826 | gi 11300389 | 668 bits (337) | 0.0 |
| CRA 117000088322182 | gi 12462132 | 563 bits (284) | 1e-157 |

Expression (Tissue Source):

| gi Number | Organ |
|-------------|-----------------|
| gi 14391046 | head_neck |
| gi 12320908 | nervous_tumor |
| gi 7312275 | head_neck |
| gi 14399233 | head_neck |
| gi 14400823 | head_neck |
| gi 14400822 | head_neck |
| gi 14377756 | colon |
| gi 11300389 | breast_normal |
| gi 12462132 | placenta_normal |

FIGURE 1B

```

1 MDLLALVLV SSLYLQAAAE FDGRWPRQIV SSIGLCRYGG RIDCCWGWAR
51 QSWGQCQPFY VLRQRIARIR CQLKAVCQPR CKHGECIGPN KCKCHPGYAG
101 KTCNQDEHIP APLDQGSEQP LFQPLDHQAT SLPSRDLNEC GLKPRPCKHR
151 CMNTYGSYKC YCLNGYMLMP DGSCSSALTC SMANCQYGCD VVKGQIRCQC
201 PSPGLQLAPD GRTCDVDDEC ATGRASCPRF RQCVNTFGSY ICKCHKGFDL
251 MYIGGKYQCH DIDECSLGQY QCSSFARCYN VRGSYKCKCK EGYQGDGLTC
301 VYIPKVMIEP SGPIHVPKGN GTILKGDGTN NNWIPDVGST WWPPKTPYIP
351 PIITNRPTSK PTTRPTPKPT PIPTPPPPP LPTELRTPLP PTTPERPTTG
401 LTTIAPAAST PPGGITVDNR VQTDPOKPRG DVFIHQPSN DLFEIFEIER
451 GVSADDEAKD DPGVLVHSCN FDHGLCGWIR EKDNLDHWEF IRDPAGGQYL
501 TVSAAKAPGG KAARLVPLG RLMHSGDLCL SFRHKVTGLH SGTLQVFVRK
551 HGAHGAALWG RGGHGWQRQT QITLRGADIK SVVFKGEKRR GHTGEIGLDD
601 VSLKKGHCSE ER (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

Prosite results:

PDOC00001 PS00001 ASN_GLYCOSYLATION

N-glycosylation site

320-323 NGTI

PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 10

| | | |
|----|---------|-----|
| 1 | 157-159 | SYK |
| 2 | 222-224 | TGR |
| 3 | 284-286 | SYK |
| 4 | 354-356 | TNR |
| 5 | 358-360 | TSK |
| 6 | 362-364 | TTR |
| 7 | 366-368 | TPK |
| 8 | 531-533 | SFR |
| 9 | 573-575 | TLR |
| 10 | 602-604 | SLK |

PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 3

| | | |
|---|---------|------|
| 1 | 213-216 | TCVD |
| 2 | 392-395 | TTPE |
| 3 | 453-456 | SADD |

PDOC00007 PS00007 TYR_PHOSPHO_SITE

Tyrosine kinase phosphorylation site

246-252 KGFDLMY

PDOC00008 PS00008 MYRISTYL

N-myristoylation site

Number of matches: 10

| | | |
|----|---------|--------|
| 1 | 156-161 | GSYKCY |
| 2 | 172-177 | GSCSSA |
| 3 | 194-199 | GQIRCQ |
| 4 | 238-243 | GSYICK |
| 5 | 268-273 | GQYQCS |
| 6 | 283-288 | GSYKCK |
| 7 | 497-502 | GQYLTV |
| 8 | 509-514 | GGKAAR |
| 9 | 538-543 | GLHSGT |
| 10 | 552-557 | GAHGAA |

FIGURE 2A

PDOC00010 PS00010 ASX_HYDROXYL
 Aspartic acid and asparagine hydroxylation site
 Number of matches: 3

| | | |
|---|---------|--------------|
| 1 | 151-162 | CMNTYGSYKCYC |
| 2 | 233-244 | CVNTFGSYICKC |
| 3 | 278-289 | CYNVRGSYKCKC |

PDOC00016 PS00016 RGD
 Cell attachment sequence
 429-431 RGD

PDOC00021 PS00022 EGF_1
 EGF-like domain signature 1
 92-103 CKCHPGYAGKTC

PDOC00021 PS01186 EGF_2
 EGF-like domain signature 2
 Number of matches: 3

| | | |
|---|---------|-----------------|
| 1 | 92-103 | CKCHPGYAGKTC |
| 2 | 160-174 | CYCLNGYMLMPDGSC |
| 3 | 287-300 | CKCKEGYQGDGLTC |

PDOC00913 PS01187 EGF_CA
 Calcium-binding EGF-like domain signature
 Number of matches: 3

| | | |
|---|---------|-----------------------------|
| 1 | 136-160 | DLNECGLKPRPCKHRCMNTYGSYKC |
| 2 | 216-242 | DVDECATGRASCPFRQCVENTFGSYIC |
| 3 | 261-287 | DIDECSLGQYQCSSFARCYNVRGSYKC |

Membrane spanning structure and domains:

| Helix | Begin | End | Score | Certainty |
|-------|-------|-----|-------|-----------|
| 1 | 3 | 23 | 0.942 | Putative |
| 2 | 163 | 183 | 0.765 | Putative |
| 3 | 397 | 417 | 1.209 | Certain |

SignalP results:

< Is the sequence a signal peptide?

| # | Measure | Position | Value | Cutoff | Conclusion |
|---|---------|----------|-------|--------|------------|
| | max. C | 20 | 0.657 | 0.37 | YES |
| | max. Y | 20 | 0.713 | 0.34 | YES |
| | max. S | 6 | 0.960 | 0.88 | YES |
| | mean S | 1-19 | 0.901 | 0.48 | YES |

Most likely cleavage site between pos. 19 and 20: AAA-EF

FIGURE 2B

BLAST Alignment to Top Hit:

>CRA|50000098943299 /altid=gi|15795193 /def=ref|NP_277060.1|
nephronectin [Mus musculus] /org=Mus musculus
/taxon=10090 /div=ROD /dataset=nraa /length=609
Length = 609

Score = 1190 bits (3045), Expect = 0.0

Identities = 536/610 (87%), Positives = 569/610 (92%), Gaps = 3/610 (0%)

Query: 1 MDLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60
M LLA VL SLYLQ AA+FDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY
Sbjct: 1 MAVLLAAVLASSLYLQVAADFGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60

Query: 61 VLRQRIARIRCQLKAVCQPRCKHGECIGPNKCKCHPGYAGKTCNQDEHI-PAPLDQGSEQ 119
VLRQR+ARIRCQLKAVCQP+CKHGEC+GPNKCKCHPG+AGKTCNQDE P PLDQGSEQ
Sbjct: 61 VLRQLARIRCQLKAVCQPQCKHGECVGNKCKCHPGFAGKTCNQDESFHPTPLDQGSEQ 120

Query: 120 PLFQPLDHOATSLPSRDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALT 179
PLFQP DHQAT++PSRDLNECGLKPRPCKHRCMNT+GSYKCYCLNGYML+PDGSCSSAL+
Sbjct: 121 PLFQPPDHQATNVPSPRDLNECGLKPRPCKHRCMNTFGSYKCYCLNGYMLLPDGSCSSALS 180

Query: 180 CSMANCQYGCDVVKGQIRQCQPSGLQLAPDGRTCDVDECATGRASCPFRQCVNTFGS 239
CSMANCQYGCDVVKGQ+RCQCPSPGLQLAPDGRTCDV+DECATGR SCPRFRQCVNTFGS
Sbjct: 181 CSMANCQYGCDVVKGQVRCQCPSPGLQLAPDGRTCDVIDECATGRVSCPRFRQCVNTFGS 240

Query: 240 YICKCHKGFGLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDLGT 299
YICKCH GFDLMYIGGKYQCHDIDECSLGQ+QCSS+ARCYN+ GSYKC+C++GY+GDGL
Sbjct: 241 YICKCHTGFGLMYIGGKYQCHDIDECSLGQHQCSSYARCYNIHGSYKCQCRDGYEGDGLN 300

Query: 300 CVYIPKVMIEPSGPIHVPKNGTILKGDGTGNNWIPDVGSTWWPPKTPYIPPIITNRPTS 359
CVYIPKVMIEPSGPIH+P+ NGTI KGD G+ N IPD GST WP KTPYIPP+ITNRPTS
Sbjct: 301 CVYIPKVMIEPSGPIHMPERNGTISKDGGHANRIPDAGSTRWPLKTPYIPPVITNRPTS 360

Query: 360 KPTTRPTPKPTPIPTPPPPPLPTELRT-PLPPTTPERPTTGLTTIAPAASTPPGGITVD 418
KPTTRPTP PTP PTPPPPPPLPTE RT PLPPT PERP+T TTIAPA ST ITVD
Sbjct: 361 KPTTRPTPNPTPQPTPPPPPLPTEPRTTLPPT-PERPSTRPTTIAPATSTTTTRVITVD 419

Query: 419 NRVTQDPQKPRGDFVIFRQPSNDLFEIFEIERGVSADDEAKDDPGVLVHSCNFDHGLCGW 478
NR+QTDPQKPRGDFVIFRQP+NDLFEIFEIERGVSAD+E KDDPG+L+HSCNFDHGLCGW
Sbjct: 420 NRIQTDPQKPRGDFVIFRQPTNDLFEIFEIERGVSADDEEVKDDPGILIHSCNFDHGLCGW 479

Query: 479 IREKDNLDLHWEPIRDPAGGQYLTVSAAPGGKAARLVLPGLRLMHSGDLCLSFRHKVTG 538
IREKD+DLHWE RDPAGGQYLTVSAAPGGKAARLV LG LMHSGDLCLSFRHKVTG
Sbjct: 480 IREKDSLDLHWETARDPAGGQYLTVSAAPGGKAARLVLRGLHLMHSGDLCLSFRHKVTG 539

Query: 539 LHSGLTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSVVFKGEKRRGHTGEIGL 598
LHSGLTLQVFVRKHG HGAALWGRNGGHGWRQTQITLRGAD+KSV+FKGEKRRGHTGEIGL
Sbjct: 540 LHSGLTLQVFVRKHGTHGAALWGRNGGHGWRQTQITLRGADVKSIVFKGEKRRGHTGEIGL 599

Query: 599 DDVSLKKGHC 608
DDVSLK+G C
Sbjct: 600 DDVSLKRGRC 609 (SEQ ID NO:4)

FIGURE 2C

```

>CRA|46000103872918 /altid=gill15128103 /def=gb|AAK84391.1|AF397007_1
(AF397007) nephronectin [Mus musculus] /org=Mus musculus
/taxon=10090 /div=ROD /dataset=nraa /length=592
Length = 592

Score = 1147 bits (2934), Expect = 0.0
Identities = 520/610 (85%), Positives = 552/610 (90%), Gaps = 20/610 (3%)

Query: 1 MDFLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60
M LLA VL SSLYLQ AA+FDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQP
Sbjct: 1 MAVLLAAVLASSLYLQVAADFGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQP-- 58

Query: 61 VLRQRIARIRCQLKAVCQPRCKHGECIGPNKCKCHPGYAGKTCNQDEHI-PAPLDQGSEQ 119
VCQP+CKHGEC+GPNKCKCHPG+AGKTCNQDE P PLDQGSEQ
Sbjct: 59 -----VCQPQCKHGECVGNKCKCHPGFAGKTCNQDESFHPTPLDQGSEQ 103

Query: 120 PLFQPLDHQATSLPSRDLNECGLKPRPKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALT 179
PLFQP DHQAT++PSRDLNECGLKPRPKHRCMNT+GSYKCYCLNGYML+PDGSCSSAL+
Sbjct: 104 PLFQPPDHQATNVPSPRDLNECGLKPRPKHRCMNTFGSYKCYCLNGYMLLPDGSCSSALS 163

Query: 180 CSMANCQYGCDVVKQGQIRCQCPSPGLQLAPDGRTCDVDDECATGRASCPFRFRQCVNTFGS 239
CSMANCQYGCDVVKQGQ+RCQCPSPGLQLAPDGRTCDV+DECATGR SCPRFRQCVNTFGS
Sbjct: 164 CSMANCQYGCDVVKQGQVRCQCPSPGLQLAPDGRTCDVIDECATGRVSCPRFRQCVNTFGS 223

Query: 240 YICKCHKGFGLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLT 299
YICKCH GFDLMYIGGKYQCHDIDECSLGQ+QCSS+ARCYN+ GSYKC+C++GY+GDGL
Sbjct: 224 YICKCHTGFGLMYIGGKYQCHDIDECSLGQHQCSSYARCYNIHGSYKQCRDGYEGDGLN 283

Query: 300 CVYIPKVMIEPSGPIHVPKNGTILKGDGTGNNWIPDVGSTWWPPKTPYIPPIITNRPTS 359
CVYIPKVMIEPSGPIH+P+ NGTI KGD G+ N IPD GST WP KTPYIPP+ITNRPTS
Sbjct: 284 CVYIPKVMIEPSGPIHMPERNGTISKGDGGHANRIPDAGSTRWPLKTPYIPPVITNRPTS 343

Query: 360 KPTTRPTPKPTPIPTPPPPPLPTELRT-PLPPTTPERPTTGLTTIAPAASTPPGGITVD 418
KPTTRPTP PTP PTPPPPPPLPTE RT PLPPT PERP+T TTIAPA ST ITVD
Sbjct: 344 KPTTRPTPNPTPQPTPPPPPLPTEPRTTPLPPT-PERPSTRPTTIAPATSTTTTRVITVD 402

Query: 419 NRVTQDPQKPRGDVFIQRQPSNDLFEIFEIERGVSADDEAKDDPGVLVHSCNFDHGLCGW 478
NR+QTDPQKPRGDVFIQRQP+NDLFEIFEIERGVSAD+E KDDPG+L+HSCNFDHGLCGW
Sbjct: 403 NRIQTDPQKPRGDVFIQRQPTNDLFEIFEIERGVSADDEEVKDDPGILIHSCNFDHGLCGW 462

Query: 479 IREKDNLDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRMHSGDLCLSFRHKVTG 538
IREKD+DLHWE RDPAGGQYLTVSAAKAPGGKAARLVL LG LMHSGDLCLSFRHKVTG
Sbjct: 463 IREKDSLDLHWETARDPAGGQYLTVSAAKAPGGKAARLVLRLGLMHSGDLCLSFRHKVTG 522

Query: 539 LHSGLTQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSVVFKGEKRRGHTGEIGL 598
LHSGLTQVFVRKHG HGAALWGRNGGHGWRQTQITLRGAD+KSV+FKGEKRRGHTGEIGL
Sbjct: 523 LHSGLTQVFVRKHGTHGAALWGRNGGHGWRQTQITLRGADVKSIVFKGEKRRGHTGEIGL 582

Query: 599 DDVSLKKGHC 608
DDVSLK+G C
Sbjct: 583 DDVSLKRGRC 592 (SEQ ID NO:5)

```

FIGURE 2D

>CRA|84000042916333 /altid=gi|15430248 /def=gb|AAK96011.1|
 (AY035899) nephronectin long isoform [Mus musculus]
 /org=Mus musculus /taxon=10090 /div=ROD /dataset=nraa
 /length=578
 Length = 578

Score = 1128 bits (2885), Expect = 0.0
 Identities = 512/609 (84%), Positives = 543/609 (89%), Gaps = 32/609 (5%)

Query: 1 MDLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60
 M LLA VL SLYLQ AA+FDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY
 Sbjct: 1 MAVLLAAVLASSLYLQVAADFGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60

Query: 61 VLRQRIARIRCQLKAVCQPRCKHGECIGPNKCKCHPGYAGKTCNQDEHIPAPLDQGSEQP 120
 VLRQR+ARIRCQLKAVCQP+CKHGEC+GPNKCKCHPG+AGKTCNQ
 Sbjct: 61 VLRQLARIRCQLKAVCQPQCKHGECVGNKCKCHPGFAGKTCNQ----- 105

Query: 121 LFQPLDHQATSLPSRDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALTC 180
 DLNECGLKPRPCKHRCMNT+GSYKCYCLNGYML+PDGSCSSAL+C
 Sbjct: 106 -----DLNECGLKPRPCKHRCMNTFGSYKCYCLNGYMLLPDGSCSSALSC 150

Query: 181 SMANCQYGCDVVKGQIRQCQCPSPGLQLAPDGRTCVDVDECATGRASCPFRQCVNTFGSY 240
 SMANCQYGCDVVKGQ+RCQCPSPGLQLAPDGRTCVD+DECATGR SCPRFRQCVNTFGSY
 Sbjct: 151 SMANCQYGCDVVKGQVRCQCPSPGLQLAPDGRTCVDIDECATGRVSCPRFRQCVNTFGSY 210

Query: 241 ICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLTC 300
 ICKCH GFDLMYIGGKYQCHDIDECSLGQ+QCSS+ARCYN+ GSYKC+C++GY+GDGL C
 Sbjct: 211 ICKCHTGFDLMYIGGKYQCHDIDECSLGQHQCSSYARCYNIHGSYKCCQCRDGYEGDGLNC 270

Query: 301 VYIPKVMIEPSGPIHVPKNGTILKGDGTGNWIPDVGSTWVPPKTPYIPPIITNRPTSK 360
 VYIPKVMIEPSGPIH+P+ NGTI KGD G+ N IPD GST WP KTPYIPP+ITNRPTSK
 Sbjct: 271 VYIPKVMIEPSGPIHMPERNGTISKGDGGHANRIPDAGSTRWPLKTPYIPPVITNRPTSK 330

Query: 361 PTTRPTPKPTPIPTPPPPPLPTELRL-TPLPPTTPERPTTGLTTIAPAASTPPGGITVDN 419
 PTTRPTP PTP PTPPPPPPLTE R TPLPP TPERP+T TTIAPA ST ITVDN
 Sbjct: 331 PTTRPTPNPTPQPTPPPPPLPTEPRTTPLPP-TPERPSTRPTTIAPATSTTTTRVITVDN 389

Query: 420 RVQTDQPQKPRGDVFIQRQPSNDLFEIFEIERGVSADEAKDDPGVLVHSCNFDHGLCGWI 479
 R+QTDQPQKPRGDVFIQRQ+NDLFEIFEIERGVSADE+KDDPG+L+HSCNFDHGLCGWI
 Sbjct: 390 RIQTDQPQKPRGDVFIQRQPTNDLFEIFEIERGVSADEEVKDDPGILIHSCNFDHGLCGWI 449

Query: 480 REKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRLMHSGDLCLSFRHKVTGL 539
 REKD+DLHWE RDPAGGQYLTVSAAKAPGGKAARLVL LG LMHSGDLCLSFRHKVTGL
 Sbjct: 450 REKSDLHWETARDPAGGQYLTVSAAKAPGGKAARLVLRLGHLMHSGDLCLSFRHKVTGL 509

Query: 540 HSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSVVFKGEKRRGHTGEIGLD 599
 HSGTLQVFVRKHG HGAALWGRNGGHGWRQTQITLRGAD+KSV+FKGEKRRGHTGEIGLD
 Sbjct: 510 HSGTLQVFVRKHGTHGAALWGRNGGHGWRQTQITLRGADVKSIVFKGEKRRGHTGEIGLD 569

Query: 600 DVSLKKGHC 608
 DVSLK+G C
 Sbjct: 570 DVSLKRGRC 578 (SEQ ID NO:6)

FIGURE 2E

Hmmer search results (Pfam):

Scores for sequence family classification (score includes all domains):

| Model | Description | Score | E-value | N |
|---------|-----------------|-------|---------|---|
| CE00409 | E00409 fibulin | 108.3 | 3.8e-30 | 3 |
| PF00629 | MAM domain. | 88.2 | 1.7e-22 | 1 |
| PF00008 | EGF-like domain | 68.0 | 1.7e-17 | 5 |
| CE00283 | E00283 selectin | 13.0 | 0.014 | 2 |
| PF01278 | Ompin family | 1.9 | 6.6 | 1 |

Parsed for domains:

| Model | Domain | seq-f | seq-t | hmm-f | hmm-t | score | E-value |
|---------|--------|-------|--------|-------|--------|-------|---------|
| PF00008 | 1/5 | 77 | 103 .. | 1 | 45 {} | 17.3 | 0.0036 |
| CE00283 | 1/2 | 74 | 105 .. | 119 | 153 .. | 10.7 | 0.059 |
| PF00008 | 2/5 | 140 | 172 .. | 1 | 42 {} | 19.4 | 0.00091 |
| PF00008 | 3/5 | 180 | 214 .. | 1 | 45 {} | 19.0 | 0.0012 |
| CE00409 | 1/3 | 136 | 227 .. | 415 | 514 .. | 69.4 | 4.5e-19 |
| CE00283 | 2/2 | 226 | 248 .. | 124 | 146 .. | 8.7 | 0.2 |
| PF00008 | 4/5 | 220 | 249 .. | 1 | 41 {} | 10.8 | 0.25 |
| CE00409 | 2/3 | 232 | 273 .. | 429 | 473 .. | 35.2 | 2.3e-09 |
| CE00409 | 3/3 | 277 | 298 .. | 344 | 365 .. | 16.3 | 0.00053 |
| PF00008 | 5/5 | 265 | 300 .. | 1 | 45 {} | 31.7 | 2.9e-07 |
| PF01278 | 1/1 | 477 | 493 .. | 241 | 257 .. | 1.9 | 6.6 |
| PF00629 | 1/1 | 469 | 610 .. | 1 | 170 {} | 88.2 | 1.7e-22 |

FIGURE 2F

1 GAGAAAATTG AGATTACTAC CTGCAAGGTG TCATTACCTG GTAAGAAGCC
51 TATCAAAAGT TTGTCTCCT GAAAAAGTAG TTATTGCTAA AAGCTAGCTG
101 TTTTGATCTC ATTCTTGCTC ATTTGTTTTT AAGACTGAGA TAATGAAATG
151 TCAC'TCCCAT GGCAACTCTG CCTCTTTTTC GGAATGATCA TTGGTGGTCA
201 TAGTTGCAGC ATAATAACCA GTTAGACCTT GGAAATCCTT TAGATTCTCC
251 TTAT'TCCATG ATTTAACAAA GACTGATATA ATTAGCTACA TTTTACTGAA
301 GGGAGAAGCT AAAGTTCACA GGCAGAATTC AATTTAATCC AATCCATCTG
351 TTTNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNCT ATTGGTTAGA
451 CTAATGTCTC ACGGGACCCA GAGTGGTAGG GAGGCAATGA CAACACAATA
501 CATTAAGAGC TACCTTAGAG CATGCTAAGT GTAATAGAAG TATATAAAAC
551 ACTAGTCTCA GCCAATCAAA AAGTCAGAAA AGGCTTCCTG GGGGACTGTC
601 GGCTAAATTT AAAC'TCAAAG GGTAAGGGGA GATTGTCCAA ATGAAGAAGG
651 AAGAGGAATG AAGAGGAGGC TAAATTTAAA TGAAGAAGGA GTGTTTTCTA
701 GGCCAAAGCA AATATATGGA AAACAAAAGA AAGACGCATG TAGATGGGGC
751 ATTACACTTT TCCCTCCAGT TATTTATCCT GTTTTCATCC ACCACTCTTC
801 GTCTTTTCTC TAGATCTCCA GTTTTTTAGC CGTATATTAC CCCCTTTCTC
851 TCTAATCATC CATTGCACAC AGTGAGGTTT ATTTGTAAAT CTAACCCAGG
901 CACCAACTAA CCAACCAACA AACAAAAACA GCTAAGGAGT AGCCACTGGA
951 ACCTGGAACC ATGCCTACAC TTACAAAAAT TTGATTTTCT GCAGAAAAAA
1001 TATTTTGA'CT CCTACATTTT TGGACTTCAT TAGAAGGACC TGAATGGAT
1051 GACACCAAGC TGTTTGCCTA AAATAATGTC CCAAGCCTGA ATTGGCATGG
1101 ATCTTTCTTG AGATTA'AAAT AGAAACTTGT TTGCTAACT GAAAACAACT
1151 TAGAAATCAA AGAGCCATTT AAGTTGAAAC CAT'TATTTT CCTTTCCTTG
1201 AAGAAAATTC CTGTTTTTAC ACACACTGAA TGATCAGGAT AGTGAATCAC
1251 CCTACCACAG AACTTTCCAT TAAAAATTTG AAGTTGTAGA AACCTCAAAA
1301 AGAAAAATGAA GATGGGGGGA AAACGTTTGT AATGTAGCAA ATGANNNNNN
1351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNAGATG ATCACTGCAC
1901 AAGGCTGTCA TTCGAGGTAA GGTTGGGAAG AGGGGTGACT TGATTGGGCT
1951 CATGTACATC TTAGGATAAT CATGTTTTCA TTTGAACACA GTCATTTAGT
2001 ACAATACTTT T'ACTCTGCA ACAACTGAAC AATGATTACT TAAAATGGTT
2051 TCTAAGTGAT CCCCTCAAGA TCAGGCACCA CAACACTACT AGGCAGTAGC
2101 AGCCCTTCCT CTCTGGCAGG CACTCATCCC TAGAGGGGCA GTCCTGTCCT
2151 AATGTGCTCG GGAGTTGAGC TGATAGAGAC AGAAACAGTC TGCTCATTTT
2201 TTCTTCAAAA CCGACTGGA GAGCCAAATA TTTCTTTTGG TTAGTTACAA
2251 AATAAAAAAT ATGCCTTTGC CAACTCTGCG GGAAAAGCAA ACTCCCAGAG
2301 TTTGCCTTCG CTAATTTGTC CAAGTCGTGC TGTGGTTTTA GTAAATGCAG
2351 AC'ACTGCTCA GCTCCAGCCC CATAAACCTC TCTGCTCTAG GGCTTCTCCG
2401 CCCTCCGTGA AAAGCTACTG CTCGCCCTG CAGTCACCAC CTGTT'CGGGC
2451 GGAACCTGCG GAGCGTGCAC CTACGCCTCG GGCTCCTTTC CTCCTCCACT
2501 CCCCTTTCCT GCTGGGCACC CTGCTTTCCC TCTCCCAGAG AGGGTTTGCA
2551 ACTTTTCTCC CAGGCTGGGG CTCGCCCTGC TTGGCTAACC CCAAGAGCC
2601 ACTGCCGTCC CGCAGCGCCC CTGCCCCCGA GTTGCTGCC CCGCTGGGCC
2651 CCGGGAGGGA GCGGAGCGCG CTCACCCTTC GCCCGGGGCT GGGAGGGCGG
2701 CGAGTCGGGC GCACGCGCAC CCCCTGCCCC CCCCTGGCGC CCTCCCCGC
2751 GGGCGGTGCA GCTACCCCTG CAGCGCCTCC C'TAGCTAGA AGGGAGCGGG
2801 AGGGGGCTCC GGGCGCCGCG CAGCAGACCT GCTCCGGCCG CGCGCCTCGC
2851 CGCTGTCTC CGGGAGCGGC AGCAGTAGCC CGGGCGGCGA GGGCTGGGGG
2901 TTCCTCGAGA CTCTCAGAGG GGCGCCTCCC ATCGGCGCCC ACCACCCCAA
2951 CCTGTTCTC GCGGCCCACT GCGTGCGCC CCAGGACCCG CTGCCCAACA

FIGURE 3A

3001 TGGATTTTCT CCTGGCGCTG GTGCTGGTAT CCTCGCTCTA CCTGCAGGCG
3051 GCCGCCGAGT TCGACGGGAG GTGAGCTGGG CCCCGGGGCG CCTCTCCTC
3101 CTTCCCGCGC TAATTTTACA CTCACTGTCT TGGGTCACTT TTTCCCGCGG
3151 GGTTCCTGG TCAGAGAGGC GTCTCTCCA TCCAGAAGTT GGGCCACCGC
3201 ACAGCGTGGC GCGAGGAGAG CGGTCCAGCG GCTCCGAGTG CCCGCCGAG
3251 GCGGAGAGGG CGCGCCCTTG CGAGTCTGGG ACCCCATCCG CGGCCCCCCG
3301 AGGGCGACTC GCCCCGGCTC GGAATTAGG ACTGAGGGAG AGGAGCCGCT
3351 GGAGCTGGG ATCTCGGCTC TGAGGGCGCG GTTTAGCCAC CTACGCCGAG
3401 GTGACGCGCG AAACATCCCT TACCCGGGAA ACTCCCGCGC CTGAAGTAGA
3451 CGGCTCTTCA CTGGGGAAGC TTCCAGGCCC CGGGGGGAG GCCCGGGCTC
3501 TGCTCAGGGC TCTCGGGGCC GCTCACACAG AGAGTGGGTG CGAGTCAGCG
3551 ACTGGGCTAC GGGGAGATT TGTGGGCCTC TCCATTTGGT TTTCTTGAGG
3601 GAAGGAGACT CAAAATGAGG ACCGGAGGGT GGGCGCTCCG TGAATGTGAG
3651 CATGAGTGTG TGGATGTGTG TGTGAGAGCG CACACACTGC GCCGCTCCTC
3701 AGACTCGGGC GAGCCTGACG GCGGCGTGCT GTGACAGGTT CCAACAACCT
3751 CGGGCGCGCT CTCCGCTGTC ACTCAGCCGG TCCTCCCGCG CTCGGGGGCC
3801 GCTCCGGTGT GTGAGAGACA CTGGGTCTGT CGGGAGGGTG TGCTCGGTCC
3851 CCCTCACCTC TGTGCAATTA CAGACTAGGC TCGTCCCGGG TGCAGATGGC
3901 TGCTGCGAAA AGAGTTTTTA CTCTGGCGCA CACCGTCGCC CGGTGCGCTT
3951 CCTTCAGCGA CCTCTGCCCC CACCCCTCCG TGAGGTCCC TGCCTAGTCC
4001 TAAAGAAAGA TGCCGCACCT GTTTTACCTT TAATCTTGA ACGAATCAAT
4051 TCCACAATTG ATTGCTTTT TGTACCCGG AGCGGAACGG AGGGGAACCA
4101 GTGGAGCGCC AAAAGATGTC TAAAACAGCA AATTAAAATG TTGATACCCA
4151 AAATAGAAAG TTCGGGGTTA TTTTGTACTT TGGGCCTGGG GGTAGAAAGG
4201 CAGGTAAAAG AAAAGGGGAA TTGAAAAAGA TAAGGAGACT TTAAGAAAAC
4251 GTGATAGCAG CCAGAGTGTA GACCTTTTTT TTTTATTTTT AAAAGCAATT
4301 CTGTGCTCAC ATTTGGGTAT GTTATGCATT TCTTGACAC ATATTTGCAA
4351 CAAGAAACCC ATCACAAGAT ATGCATATGA TGTGAATGCA TATAGCTTTT
4401 TGTAACCTTA AAAAGATGTC TAAAACAGCA AATTAAAATG TTGATACCCA
4451 GAGGAAAAAA GTCATTTAGG TGTGACTCTA AACAAGGAAA CAATTTAGCA
4501 AATAATGTGT CAACGTGTAA TCAGGTGGAT AAATCTGTAT CCTGAAATTA
4551 CTTTCTTTAG GCATTATTTA CATTAGAGAG GAAAATACTA TGAATTGTTG
4601 ATCTAAGCAC GTTTCAAACC ACAAGGATAC TTGAGATATC AGCTACACTT
4651 TAACTCTCC TTGGTTTTAT TTAGACTTTT TTCTAGTACT TTTTGTGTTT
4701 TTTTGTGTTT GCCACAGTTA TGTCTCAAAA AGAGCTCTGT TACACTTGAT
4751 TTTTGAGAAA CCTTCTGCT CCTTCCCCA TCCACCTTTT ATATGAGGCA
4801 GAAATTTTTT CTGCTTTGAC ATGTCTTACT TAATACTTTT CAGTTTATGG
4851 TGAATCAGTC AAACCTGGCT TTCACCCAG TAACAGGTGT GGGTCTTTGA
4901 GGAATCTGTT TTATTTCAAG CTTCAGAAAC TATCCTGTGA GTGGCAGCCT
4951 TGAGAAATGT TGGTGTAGC TGGTGCTATT TCTACTAAG AAAGCATCTT
5001 TCCCTCCCGT TTTTCTCCT ACCCTTGTGT GTTGGGGGAT GGGGTATTTA
5051 ACTGTCTGAA ATTTGACATT AGATCATAGC AGAACTATTT CTGTGAAGGT
5101 GTTTTAACTT TAAATTTCTG CAGGCATAAA AGAGTTTGTA GAAAATGTGT
5151 GGGTGTGGTT GACATTTTTT CTGAAGTAAT AATACAGAGG AAATTACCTT
5201 TCTTCTCAA AGTGTCCTTA TTTATTTTGA ATCTTTTTTT TTGGTAGGTG
5251 GCCCAGGCAA ATAGTGTGAT CGATTGGCCT ATGTCGTTAT GGTGGGAGGA
5301 TTGACTGCTG CTGGGGCTGG GCTCGCCAGT CTGGGGACA GTGTCAGCGT
5351 GAGTATCAAG CCTGGGGACT TCAGTCCCT GGGAGGTGTG GCTTCCACC
5401 TTGTTTATGG CTTACCCCA CATATCAGAG GGTTCATTAC TGAGCAAGGC
5451 TTGGCCTTGC AGGTCTGACT TGGGGATTTT CAGGTACAGT CCAGACTCCT
5501 TATTCTGCTT CTTTTCAGCT TTAGCCACCT GTATTACGGC CCAGCTTTGT
5551 CATTCACAGA GAGCCACCTT AAATGTTTCT TTTAGCATTT TTTCCCCCAA
5601 ATCTGCATCT CCTTCTGTT TTCTTACCAG GTTTATTNNN NNNNNNNNNN
5651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5751 AGATGGGGAT AAGATAATGA GTCTAGACCG TATTTACTACT GGGGATAAGA
5801 TAATGAGTCT AGACCGTATT TACACAAAAT AAATCTGAA CCTTAGAGTA
5851 CTTATATAAT CCTCACCATA GCTGTGTGT AAAGTGGCCC ATTTGTAAAT
5901 CTTTTTTTGG ACTGGGGTTG TATTGACAGT GTTTTAAATG GACCACCCAG
5951 TATAATGAAA CAAAGCCACT GAAACAACAT TATTTGTAAG TTCTATAAGT

FIGURE 3B

6001 AACCAACCTC ATGTAATAAG TACCAAATCA GAAAGCTGAT TTTACCTTT
6051 TCTGTGACCA CAAATGGATA TTTAATGGT AAAATTTAGA GCTACATCAA
6101 AAGAGTTTGA GAAATTTAGG AAATTTGTCC CCAGCATTTT ATCGCGATCT
6151 TAAAATTGTA TCTCACTCCT ACTGCAAAAA AATAGTCTTC AAATGATCAA
6201 GTACCTTTCC AGAGCACCCCT TAGAGTGCTT GGGGCGGGG TAGGGGGCAC
6251 TTCTGGTAAG ATGATGGGAA CTAAGTTGGG TTCTACATTG GGATATATAT
6301 TTTATTGCTA ATGAGGAGGA GGCTTAGAGG AAGAGAGAAG GGCAGTTACG
6351 AAGGCTAGAG CTGGCAATGG AGAAGCCTGC CTTAGAGATG GGTGCTAGT
6401 GTGAGGAGTC AGGCAAATTT AAGTTCAGGA AAGTTAGGAG TTCCCTCTGC
6451 TATTTTAATT TTTGAGGATG CTTGCAATGT CTTCTTAAAT TTTGTGAAAG
6501 AGGGACAGTG ACAGTCACAG ATTGACTCTA ATTGCACATA AAGACCACAA
6551 TCTCTGGTTG GGAATAGAAA GGTAAAGGAA ATGAATGTTT GCCTACCTGG
6601 TATGGAATTT GAGAACCAAC AGATTCTAAT AACCAAAATG TGAAGAAAGG
6651 ACCCTTCTGT TGGCCCAACA CACCTACACA TAACCCTCCT GAGTGAAAAA
6701 TGAGTAGTTC TATACCTGCA GTCTCCAGCT GTGCAAATAC TTCTGATACT
6751 ACAGAAGACT AAATCCACC AGGCACCATT CTTCTTTTTT GATCATCTTC
6801 CCTTAAACAA ATATTGAATA GACTAACCAG TGAGTGTACA GCAGCTTTGC
6851 CTACTTCTTT TGTTTACTGG AAAGTGGAGT TACCACCATC TCCCTTTAAC
6901 AGAATGTAAT TGACCCCCCG TGCAAAGGGT CTAGCCAAGC AGCTCGCACA
6951 TAGCAGTTAG GCAAACATTC GTTTGCTTCA CCTTTTATGT TTATACTTCC
7001 AAAATCAAAG TAGTTCCAAG GTTTTCTATG TTACAGTGAA ATAAAATCCC
7051 TTTCAATTAA AAAGGCACAA ATGGTTCTTT ACTATTAACA TTGAAGTTGG
7101 TGAGGTTGTA AAAATTAGCT CAAAGGTGAA TGTTCCTTC TGTGTATTTT
7151 ATTTTCCAG CATCTTAGAT GGCACAAATG TCTCTTGTC GTTCAGAGTT
7201 CTGCCTGTCT TTGTTTTGAT ATAAGCAGGT AGAGGAATGT GGGGCTGAGA
7251 AGTAAGCTTG AAGGGGCAGA ACAAACCAAA AAGAGGCTGA TCAGATTGAA
7301 TGAAATATCT CTGAAAACATC TTGATTATTT TAAAGAAAGT CTTTATGAAA
7351 TTAAGATTTT TGTCACTACG TTTGTTCAAG AAAATGCCTT GCTATTGTAT
7401 AACCAATCAA TCTAATATGA TTCCTTATAA AGATTCCAAA GAACCTCTAG
7451 CGAATTTAAT GTGAGAAATG TTTTTCCTT TTCGACCTTT AGATAATCAT
7501 GTAGTTCTTT CCCATAAGGA AGGGCTATTC TCCCTTCCTC ATCAGAGGTG
7551 CTCTGGTTCT CTCTTTTCT GAATGATTCA CTTTGGAATT TTCCTTCAAA
7601 ACAGCATAGC AAAACAAAAA GAAACTATTC CCATTACTGC ATAGATCTTC
7651 CCAACTTATC CCATAGAAGG TGTGTTTGTA GGTAGGGAAG GGTGGTGGTA
7701 CCCTCATTAT ATTATTTAAC AGACCTTATG CCTTAGTGCA GTGACCTTTT
7751 GGTGAGTGTT TATTCCGTTT GGTTGAGCAT TTGTGAGATG TTTGACTTAA
7801 TATTCATGTG AGTCAAATGT ACGTATCTGT ACAAACATG CTGCCCTTCA
7851 TTTTCACTTG CTATCTTCCT ATCCATGGTC TTGTTTGGAG AAACCGACTA
7901 ATGTTGCAGG ATGCTAAAGC TGGTAGACCT CTCCTTCTGG CTCACTATGT
7951 CTAAGCAGAG CCAGATATAG CTGGGAAACT TTATATCCTT CCCTCTAGGA
8001 CTCAGGAAGG AAGGATCAGA GATGCTACTC AAATGGGCAT AGAACCTGTC
8051 CTGCTGCTTC CTGCTGTAC CACCAACAT TCTATCCCAA CATTCTGCC
8101 TGCTAGGGAA AGCATGACCC GTTCCAGAAA TAGGCAGGTT GTGTTTTCAT
8151 AGTTCTTTTG TAGGTCATAC CCTTGGTTGN NNNNNNNNNN NNNNNNNNNN
8201 ATTCTTTGTT TTGTGTTCTT GACTCTTAAA CAGATCTCTA GCATATTGAA
8251 AATTCAACAT TTGATTTTCT AACTGTCATG GGCTTTACTT TTATTGACTA
8301 TTGATGTGGC TGTATTATTG CAGGTGAAAA TTTTAAATA TGTTACACA
8351 CTGATGATTG CATATTTGCA GCACACAGCA TCTTAAACCA CTCAGAGGTT
8401 TGTACAAAAA TGTGTGTTTC TTGCTCTGTA ATTTTTTTGT CATTTTGATG
8451 GCATATTTTA ATTATGCTTT TATTTCTCTC CTTCTACCAA GTGGAGACCT
8501 CTGTAACAAA GATTTTTCAG GGTGTCATAT TTCATCTTTA AAAGCTGTGT
8551 AACTGTGGA AACAATTGTT AACCAAAAGC ACTAGATAAT TTAGGATAAC
8601 TGAGTTCAAG TTCTGACTTT TGTACCTGAA TAGGTGTCTA ACATTTCTCA
8651 GCTACAGTTT TCTTATTACA CAAGCATATT TCCAAGGTAT GTTGAAGCTC
8701 TAAACTGCAA CTTGAAACTT CCTTAATGTA GAAAACTAT ATAGGATCTA
8751 AATATTGTAT TTTTGTATGG TTGTCCTTCT GTTTACTCGG AGATTTGACT
8801 GTATATGTGG CTTATGACAA TAGCATTTTT GTTAAAAGCA TTTTATAGAA
8851 GTGTGAAGAA AAATAAAAAA TACAACCAGT TCCAAGGTTT AACAAAATA
8901 TTCCGTTTCT GAGTTCTTTG GCTGTCATTG AGCAACTTGT GGTTCCTGAA
8951 GGAAATTATG TGAATTAGGA TGGTTTTGTA TCATTTATCC TTAAGAACAG

FIGURE 3C

9001 GGAAAATTGA GATGTTTTCT TATGTTTTCTG CTGGAGATTT TGGAAAGATG
9051 TGAAACCTAC ACCTACAGAT TGACCTTGCT TAGTTAGCTC TGAACCTCCT
9101 GCTGCCTCTT CCACGTAAAG TGAAAATTTT GGATTCTTAT CGGCTTCAGA
9151 TAAACTTACA GGTTAGTGAA ACATAGGGAC TGAGATATAG TAATTCATTC
9201 TGAAGCTGTT TTGGAGTGGT CAAATAATTT TAGTTGGATA ATATATATTA
9251 CTGGCTAATG ATTGTGGATA TTGGAAGTGA TGAAAAAATT ATTGAATTAT
9301 TTCTTTCTGC ATTTCAAATG AAAAGGCTAT TAGTTTGAGC AGAGAATTTT
9351 GATTTAGTAA ACAAATATT TAAATTTTCAT GTTTCATTTT TTTCTCCTAT
9401 CTGGGTTTAC ATACTCAGTC TTATAAATGG AACATGATTT ATTTTGTGCTC
9451 CCTAAACTGG TTATTAACCT CCTGTCCATA ATCACAACAT TATATAGATT
9501 ATATATTTCT TTGATTATTT GGATTTTGAA TACTCTCTTT AAAATAATCA
9551 AGAGAAAATT AGAGCTGTTA GAATGTTAGA ATTTGTTTTG AAGGCCACAC
9601 ATAGTGTTCC CTCCACAGAG AGGTTATACT AGTAAATGCC TTTCTATTTG
9651 AGGTCAACAA CTATGACAAC TTCCATTGAA CATGAGTTAG TATTTTAAAC
9701 GTAAAGCAAT TTTTATACCT GTATGCACCC AAAAAGTAAC AGGGGCTCTC
9751 AAAAGGGGTG GGGATTGTAC TGTTTACATG TATATTGAAG ATTGCTAGCA
9801 GAATTCCTGGG GCCAGCTTGG TGGAGCGGAG TACACTTCAT TGTCTTAGT
9851 GTAGTAGCTT CCTCCTCTAA TTTTGAGGTG AGAATGCAGA ATCTGTTTTT
9901 TGTTTGTGTTG CTTTTAGGTA GGAATAAAAG CAAGGCAAGG AATAATTTTG
9951 ATTACTTGCA ACATTAAACT TGAATCCACA AATCCTTAGG AAGTGAAGTT
10001 TTTGATTAAG ATTATTTAAC TGCCACTTTC CTTGAAAGGT TGTTTAAGAA
10051 CATCATGTAC CTTTGGGTAA CTCAAGTGG TCTTGGAATG CAGATTCCAA
10101 AGTAAGATCA CGGTGGAAGA ATCTTGACCT TTTCAAACAG GTAATTTGTT
10151 AGTATGTGTA GTCTTCAAAG TTAAGTTTCA GAGAAATTGC TTTGCTCTTT
10201 TATTCCTTTT CCATCAGAGA ACTGATCTTT ATGTAATACT ATTAAGATCT
10251 ACTAATTTCC TGAAATCCCT TCATAAGCTT AATCTGGCCA GGTCTTAACC
10301 TTTATGGATT AGAAATTTTA GTACTTCTTA AGCTAGAAGG CCAGGCCAAC
10351 TAAAGGGAAG CACTCTCCTC TCTGCCAGTT CAGCAACTAG ATCTGTCCTC
10401 AAATACCTGC CACAGGGATA TGCTGCTTGA AGTTGCCAC GCATGTACAC
10451 ACTGGGACCA AGAAGGCACT TCTGGTGCCA GAAACAACAC TGTGTTTGCT
10501 TTGTGGAAAT TTTTGATATG CTTTAAAAA TGTAGGTGTC TTCTCCCTCA
10551 CCTCCTGTGT GTAACCTCCA GCATTTCTTT TGTTTGCTTT TTATCTACAG
10601 AATTCATGTT CTTTGCATTT TGAGTTAGTT GAATCCACTG TATGCTTTCC
10651 AGATGATAAT TAGTGAAGCT CAATGATTCT ATGCAGTGC TTTTCTGCTG
10701 TAAGGGAAGA TATTCATCAC AGGGGTGGGC CTTTATCACA GGGGAGATGG
10751 AAGGGCAAGG GAGAGGGATT CTGATGGCCT AACCCTCATT ATTGCCCCC
10801 ACTGCCACAT GCAGACACTC TTGCCGCTTT CCCATTGATT TTGTGCAATA
10851 TTTTAAATTT TTCCACCAAT CTTTGGAAA TATTAATTTA TTTTGTAGGA
10901 GAAAATGTAT TCATTTCTTT AAATATCCCT CCTCAGAGCA ACAAGCATGA
10951 TTAGTTTTTT GTATACTTTA AAAAATATTT TACCTACTTA TGGGTAAATT
11001 GCAATAGATC TCTCCCCTAT CACCCACTTG TTTTGTACAA ATGAAAGCAT
11051 ACTATATACA CTGCTTTGTA TGCTGCTGCT TTTTCTTAAC TGTACATGTT
11101 GAAAGGTTTT ATTTATAGAC CTCTTACNN NNNNNNNNNN NNNNNNNNNN
11151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNC CTCCTTACTT TTTTATGCAC
11451 TCNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
11551 AAAAAACTTT GAACATATGT TAAAGATGTT GTTTTGCATG TTCTTAGTGT
11601 TTCTGCAAAA TAAATTTTTA TATGTGGTCA AATGGCAGAT ACATTTGTAA
11651 TTTTATCAA TTGCCCAATG ACTTCTTACT CAGTAGAGAA GATAGTTTAA
11701 GGGGAAAGAC TGAATGGCGA TGCTGACCCA AAGGCTCCAA CAATCCCAT
11751 TCCAGAGTCT GATCTTTCCT TGATTTTAGG ATTCTTTGAT CCCTTTTCTT
11801 TCCCAAGAAA TCCCTCTGAC AACTCAGTGA ATGTCCATCT GCTCCCATCA
11851 TCTGTTTCCC ATCAGACCAG AGGCAATGGG CTGAAGCTAG GAGAATGGAG
11901 TGAGATGTGT CCATTTGCCA GGGTCTCCCA GTGGCTTCCT GCCTATCATT
11951 TACTTGTAAG AAAGAGCCAC GCATTCCTTT AGGAATTGCT TAATTCATTG

FIGURE 3D

```

12001 ATCTTTTATT ATCTTTGTAG TTTAGCCATA ANNNNNNNNN NNNNNNNNNN
12051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12201 NNTCAACTGA TGGAGTTGAT ATCCAGAGAC AAGATACAGA GATTCTGTAT
12251 CTTGNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNTAGCCATA ATTTTATAC
12401 CTTTGTGTAT AGTATGGGCT TTTTNTTTT CTAGACTTAT AACCTATTAA
12451 GGAGGAATAT GTAGAATTCT GAGATCTGAA TACCATTTTG GGGACACCGT
12501 TTATGATAAA AAGATTTTCA GGAAATATTA AGACATTTT GTACCAAAT
12551 ACTTTTGTGA GTGTCCTTAA AACCATAGGG GCATATATTT TCTAATAGTT
12601 ATTCTAATCC TTATCCAAT TATATACTTA AATTGTCTCC TTTTAGAAAA
12651 GTAAGAAATG CATGATTAGG TGAAAAATTA AAAGACACAT ACTAAGAGCA
12701 AGACAGCAGT GAAGGTAAGT ATTTCTAACT TTCCTGTTTT CCTGACACCT
12751 AGTTNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNCA TCATCACTGT TGTATACCTT
13101 TAGATATGGA GGACATATCT AAAGACATTT GATTTCTGTA TTAGTATAGT
13151 AATTGAAGAG GTTTCATTTT ATTGCTACAA ATTTTCTTAT GTTTAATCTG
13201 TAATGTAAAG GAGAGTAAAT CATGGCTAAT ATTACAGAGT AGAAACCTTC
13251 TGTTCAAGTCA TCTTCTATTT TTGATTCCTT AGAAGATACC TTTAAAGATA
13301 AATTAAATTG ATTTCTTTTAT TTCTATTTGT TTATGCCCTA TCTTGTTTGG
13351 GAAAGTATAT AAAGTGGCTT CCAGGAATTC ACACAATTAA AATAAGGAAA
13401 TAGGGGCTAC ATGGAGGAAA ATGGGGGTAA AATAAAAATA TTAGGAGGAG
13451 GGTTTTTCAA ATGCAGATAT GTAAGCCATA GGATCCTACA TATGTCCTAC
13501 AATTTGGCTA CTAGCTTCTG GTAGCCAAAG GTAAAAGATG TTATTGCTGA
13551 TTTAATTTGT GTTGTTTGTC AGGGAAAAAG CAAACAAGTT ATGTGGGAGG
13601 AATTATGGTT TTAGTGCATG TAGTTCTGAA CGAAGACTTT CACGTGGCCC
13651 TTACGGAGT GGGAGTGAAT GATGGGATGG AAGGGGCCAC AGCAACATTC
13701 CTATCATGAA GCCAGTGGTG GTTTTGAAG AGCTGTTTCA TAGAAGATCC
13751 GTCAACATAA GATGAGGCCC TTAGTCTAAA AGTACAGTTC ACGGAAAGCT
13801 GTTCTATACA TAGTTTAAAG AGCATTATGT GCATAGCTTT CTGACGTTCA
13851 GCTAGATACA GCCCAGAATA TCTAGCTGTG ATGGATGAAC TGTATAGGGT
13901 TTTGGATGGA CAGCTCTGGA TTAGACAAGA TAGTTTCAGG TTAGAATCAC
13951 TAACAATGTT CTGAAGTTTG CTGTATTATT TAACAGATTA AAGGCCAGTT
14001 CATTTTGTCT TTCTTTTSTA AGTCGATATA TTTTGAAGAT CAGTAATCAA
14051 CTAATGGAGT TGATATCCAG AGACATTGTG AATAGAGCAG GGATAGGGCC
14101 CCGCCTAGGC TCCAGTAAAA GGAGGATCTG CAAAGAAAT AGTGGGAGAT
14151 TTTTAAATAT CAATTTTGTA TCCATTGGAT AGCAATATAT AAATTACCA
14201 CCATTTTGCT ATATACATAT TCATTTCCAT GCCAGTTATG GCTTGGAATA
14251 AGGAGGAAAG GCATGAACAT TGTTCCAGG CATCATTTAC CCTACTGATG
14301 CTGATACATG GATATGGGCC AGTGAGCCCA GGAACAACAG AATCCCTGAC
14351 CAGTGCTTAT TTCCCCAAG TTCCTGATTG CTATAGGTGT TCCGGGGAGT
14401 CAAAGTAACC TCAACTGTTT TCTTAATTCA CCTGGATTTA TAAGGGTTGA
14451 TTTATATAAG TTGCTTAGAC CTGAACAGAC TCAAAGCAGA GTCTGTAGGA
14501 AATACTCTGC ATATCAACAT CCCGTACCCA AACCTAAGTC ATCTTTTCAC
14551 TGGGGTGTGC GGAAGGGCTG ATTCTCATGT ACTCTTGAAG CCCTAGGCAA
14601 TAGAACCTGA AATCCTGATG CACATATACC CTAGGATAAT TTCTCTCTCA
14651 AAAAAAGCA AATAGTGATT TTAATAATTA TGACTCAATG CATATTCTGA
14701 GAGTGCAGCA ATCCAGTAAG TGTACATCT CCTTGGAGG AACAAAGAGA
14751 AGGTCTTCAA ACGCCTTTTC TGCAGAGACT AGCGTGAACC AAGAATCTCC
14801 CTGTCTGAAT TGTCACCGTA TATCAACGTG GCAGCGAGCT GAGAAAGTTA
14851 GCATTGCCAG GGCCGAGGTT TTCTGTCTTC ACGGTAACAA ACAAATGATC
14901 AGTCTCAAAA GATAGTAAAA TGTAGCCAAG ATAATTTGGC TTGAAAAAAA
14951 TTCAAAGTGA TGTCATAAAC CAGCTACAAC CACAGAAAAA TACTAGTAGC

```

FIGURE 3E

15001 ATATTGATGG TGCATCCCTG GGAACATGCC AGGGCAAAT T GGGGTGCAAA
15051 ATACACCAGT TATTTAAAAA TTGGTATGTT ATAAGGTAAA GAATTCATTA
15101 AGTAATATCA AATACAAATA AACTTTTTAT TTTGATTCAC TAAAACTTCT
15151 TTTTAAGTTT TCTGATTTTA TTACTTAATT ACTACTGATA CATAATTAAA
15201 GAATTATGGC CATAATAGAA TTCCTAATAA AATTTCTAAT AAAGCCAGAA
15251 TGAAGGAGTA TATTACAGGC CAGACATGAT AAAGCATTAT GATGTGTGGT
15301 AAAATAGTGA CATCTATTTT TTCATTCTGT ATTATTTTAT AAATTTTCTG
15351 GAGAATTTCA GTTTAAACAG CCTGCTGAAA TACTGTTAAA TCAACCTGTT
15401 ATTCTTAAC TGTATGGGAG GAAACAGATA TGAATAATAA AAATGATTTT
15451 TTAGCTTTAG AATATAGTTG TTGCTTTGGA GAACAAACCG TTTTATTCCA
15501 ATTATTTTTT ATTAGGAGAC TTCATTTTCT GCCATACATT AGCTTTGGTA
15551 GATACTAGAT GCCAGGGGAG TGCAAATTTG AGAATAAGGA TTTGGCATGG
15601 GTTATTTGCA TGTGAGAGT CAAGATTTAA CTAATTTTAA AAACCAACA
15651 CTTCAATTGA TCAATATCTT TTTAACCATT CTGTAGATTA ATAATATAAA
15701 TTCTCCAGAG CTGACATTAC TTTGCTATAA CATCATCAGA TCACAAGATT
15751 AGGGTTGCCT TTTGTAGATG TTATTCACCT ACTCTGAAAT GTTAGGAAAT
15801 ATGTCATCAC AGTTTAAATT TGTAAGTAATA TATACAAAAG GAAAAACACAA
15851 CTAGGAATTT TGGATTTATG CTTACTTTGC CAAAAACCAT GTTGATTTTC
15901 AAAAACCTTT AGCCNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16201 NNNNNNCTAA CCTTATTTTC GTAAATCAAA GTTCCTAAAT TCAGTTTCTA
16251 AGGAAAGCCC CTAAGTGGTC AAAGGTGTGT GTTGCTGAAT TTCAGAAACA
16301 AACTCCATGG ATGGACTTTC CAGACATAAA CTTCCCACCA AAATATTGTT
16351 CCAAAGTGT TAGTGCCTCC TATTTTATTT TAGTGTTAAG ATTTTGTAGG
16401 TACTTGGTAA TTATCAGCAG AATATTTACA TACCAACAAA TAGCAAAGCC
16451 ATTGATAATT AATAGAAAAA CAAAATCATA GCTTATGTTT AGTTGTATTT
16501 TTTAGTGTAT TTCAGTGTTC AGAGTTGATT TTATATTGCA TATACTAACT
16551 GTGTAATAG TTTCTACTGG ATAGATTATT TCAAGTTCAC ATGATTATAG
16601 AAGTATTTTA TGAAGTTTGA CATAGATAAT GGTGGTTTCA GGGTCCCTTG
16651 TCTCTGGGT TGAGTTTFTA CATCATTGAA GAATTAAGG ATTTGTGTCTT
16701 GGTGTCTATG TATGGCCAAG AATTAAGACA TTCTGTGTA CAGTTTCATT
16751 TTCTAACTAA GGCAAAAATG ATTTTGTAAA CAATTACTTT TTAGGATAAA
16801 AGTATAAAAT TATGAAGAAA TAATGTGTAG GTTTGCTAAA TGTGTGTGTT
16851 TTAAATCTTG TGTGAGGTC TGATGAATTT TTTTATATCT TTAATAGTTC
16901 TAAAGTAGTG TACTGAAAGT TAGGATCATC CATGTTGCTA CTTAACTTGT
16951 TGTCATAGG GTCTGTTGTC CTGGATGTTG GTGTTATAAA TGTACAGAGT
17001 TAGGTAGTTT TCTGTGAAGA GTTTGCAGGC TAATATCACT GTTTTGTGACA
17051 ATGAAAAATA TAGCATGAAA ATTAAGGTTG GGGTATGGAG AACTTTCTAG
17101 AGCTATTGCT TTTCTAGCT GATGATTAAG GTTGAGAGGC AGAAGTACAT
17151 TGTGTACAGG AGGACCTTTC CCTTTTATGC ATGCACCTGG CATTTATCAC
17201 CTTTAACAAA GTGTGTGTGT CATGCTGTGC TGCTTGCTTA AGGGGCTGTA
17251 TGCTCCAGT CTGAGTCACA TGGTAACTGC ATCAGCAGTC TTAGCCTGTA
17301 GCATTTTATT ATTTCTTTTC AAAGTTTACA CTTGGCCTAA TACTTAGACA
17351 TTTTATAATT CTTACAGTA AATGAACTA TACATGAGAA GATGGGCATA
17401 CCTTTGGATT AAAAAAATAA AAGGCTCAGT TCTTTAAAGT TTCTTATCCT
17451 TGATTTTCTT AACACGGTCC AAAGTTCAGG ACTGGCTCCA AACCATAAAA
17501 CCTGTGTTTA GCAAGCAAGA AACATAATCC CCAGATAGTT TAGTTTCTCT
17551 GGTTTATGAG TCACAATTTT ATAAAATCAT CAGAGTGCTT ATTAATTCCA
17601 ACCACGTATA GTAAAGAACC TTCAGATGAA CTGAAGCAAG GGTTCCTTAGG
17651 CAAGTTGCCG GAGAGAGTTC TAGAATTCTA AACTACCTGA GTAGCTTTGC
17701 TGAATGTTG CTTGTATTG CTACTGTGGC CATTTTATGA TGGCCATAGA
17751 GCAACAGATT ATCAAGAGAA AATGAGACAG ATTTTCTGTT ATTATGTGCT
17801 ATGAATGAAT CCTATTTTAG TGAATGTTTT AATGGGGTTT ATACCGCAAA
17851 AAAAAAAT TGATGTAAAG GCATTTATTT CTGGCAACTT TCATAAAAAT
17901 TGTTGTGATG GTTGCATATA AAAATTTTCT TATCCTTCAA TAGAGGATAG
17951 TTCCCAAGAA CTTCTAGAAA GTAATCTATT CCAGATTTAA CATTGCTTTG

FIGURE 3F

| | | | | | |
|-------|-------------|------------|-------------|------------|-------------|
| 18001 | ACATAAAATG | CAGTTTTGTG | TAGTTTTTAA | ATGCAAATTA | AAAAATATAG |
| 18051 | GACATTGGCT | AAAAATTTAT | CTTGAAGTCG | GGTATATATT | GATACCATAA |
| 18101 | AACTTACTAG | ATCTATGTAT | TTCAAGGCTA | ATTTATGCCA | AGTAGGAAAA |
| 18151 | ATATGACCCA | ACCTTAAGAT | ATTACAAGGA | TAAAATAGAC | TATACAAAAC |
| 18201 | TGTTTGGCTA | TTTGGTACTA | ATACAAC TAG | TTAGAACATA | ATGATGTTTG |
| 18251 | CTATTCTTTA | TTAAGTTGTT | TTACCTTTGC | TTACAATAAT | TTAAAGTATT |
| 18301 | TTTCCTGATA | AATTTGATGA | CTCAAAATTG | GCAATTAAAG | AATATTTAAAG |
| 18351 | AAACGGTATC | CTTTTATATT | TTTTCTGTCT | CATATATAAC | CATAGTCATA |
| 18401 | ACTTGTGTGA | TCCAGAATGT | AATTTGCTAT | TTACACTTTG | ATTTTCAGCTG |
| 18451 | TTTGCTTAGA | TTGTACCTGA | TGTATTTTAT | TATTCTATTT | AAGGAATGTG |
| 18501 | TCAACATCAA | GATATGGTGA | GTTCTTTTCA | AAATAACAGT | AGAAACCTGA |
| 18551 | CCAATATGAA | AAAAAAAAGT | CTAGCAAGCA | AATGTAATTT | GTGTCTTTTAA |
| 18601 | AAATACATAG | CAATCATTCT | GGATCAATAG | TTAAATTATT | GCTTCAATTA |
| 18651 | AATCAAATTG | GACTTAGAAT | TTTTTTCTTC | TTATATTACC | CAAAGGAAGG |
| 18701 | CCCCATTAC | CNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNAA |
| 18751 | TTGTAAATTA | CAGTGGATAT | TTAATCCTTT | AAAGGCATTA | ATTTAGCAGA |
| 18801 | GAAGAATAAA | ATTATCCCTT | TTCTTACCTC | TAAAATCTCT | AGGTTGATCA |
| 18851 | AACACTGACC | ATTATTAGTA | CCATTTCAAG | TTTCTTCCCT | TTACTTTTCAT |
| 18901 | CAAAC TGGGT | TATTTATGTA | TGTTCAAGTG | AAATGGCTGT | GCTTTCATGG |
| 18951 | TAATTCTTGT | TGCTATTGAC | AACCAAAGCA | GCCATGCAAG | AAGAAAATGC |
| 19001 | TGTGGAAGGG | AAGAAAAAAA | TTATATTTCC | TCCCCAAAGT | TGGAGAGAGA |
| 19051 | AGGGAACATA | CAGTATGTAA | GAAACAGGTA | AGTAAAAATA | TATCAATTTT |
| 19101 | AAAAC TATTA | GCCTTTCTTA | CTAATTGATC | AAAAGTTTAA | ATTTTTTTCAG |
| 19151 | ATGTGTTTTT | CAACTCCAAT | TTAAATAAAA | GGATACTGTC | TCTAGGAATC |
| 19201 | AGAAATTAAT | ATTCTGGAAA | TGAAAAATTG | GATTTGAAAA | TACAGCATCA |
| 19251 | CAAAGGTCT | GAAATATTTA | AATTTAGAAT | TTGGACTATA | GAAGAATAGC |
| 19301 | ATAACTTCAA | TTATTGGTTA | ATATTTTTGT | TATGAAATGT | TTTTTTTATAA |
| 19351 | CAAAATGGTG | ACTAAAATAT | TACTATTTTA | AACATGTCCT | AGATTTTTTT |
| 19401 | TTGTTTCAGA | AAAGCACTGA | AAGTTGAATA | TGTGTAAGTC | TCCGGGAATG |
| 19451 | TAACAAGTTG | ATAAATACCC | AAGTCCAGCC | TCCTTACAGA | GAAAAGATCT |
| 19501 | GGAATTTCTT | TTTTTGGACA | TTTGTTTTGC | AGCTTTCTAC | GTCTTAAGGC |
| 19551 | AGAGAATAGC | CAGGATAAGG | TGCCAGCTCA | AAGGTTAGAT | GAACATATTT |
| 19601 | CTTGAAATAA | TTTTGGCTAA | TCTATGTCTT | GAAAGGCATA | CCTTCTAAAT |
| 19651 | AATTTTACAA | AGCTGTAAAC | AAAACATTAG | TTGTGTTTTT | GAATTGCTTC |
| 19701 | TTTTTAGGAA | AACTATATTT | CTTAGAGATG | TGCTTATTCT | ATACATAATT |
| 19751 | ATAGTAAGTT | AATTTTAATT | CCATTATTAC | ATTAAACTTT | TCCGTTTCAG |
| 19801 | TTGATTATTC | CTTCTTCTG | TTTTATGAAT | CTAATAAATT | GACCTCCATT |
| 19851 | CTAAGCTAAA | GGACAGTTTA | TGTTTTTAAT | ACTTAAGGAA | GCAAATATAA |
| 19901 | AAACAAAGTC | CACCAAATGA | AATTACTAAC | ACCAGCCTAT | CATTAACTTG |
| 19951 | AGGGAGTTGT | ATTTCATATT | CATTAATTCT | CCATGGAACA | ATTCAGGTAA |
| 20001 | GAATATATTT | TATTGTATGA | AATATGTTAT | CAGTTATAAT | AGCACTTACG |
| 20051 | GGCATTTATG | CTCTGTGTTT | AGGTACATAA | AGGGTTCCTT | GTTTAAGATA |
| 20101 | TTTTTGGCAG | TAAATAATGT | CACTTATCAC | CCTATGACAC | TTTAGTGGAA |
| 20151 | AACGTTGGTG | AAATATGTGG | AATTGTAATG | TTTAATGGAA | TTGATTCTAA |
| 20201 | GAAGACAATG | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20451 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20501 | NNNNNNNNNN | NNGAGAATGT | TCTTTAAACT | GATTGAAGAA | ACTATCACAT |
| 20551 | GTCA TGAAAA | ATGTTTTTAA | GGCAGNNNNN | NNNNNNNNNN | NNNNNNNNNG |
| 20601 | AACAAAATTG | ATAGTACTAG | AAATAGAAGT | TCAGAAATTT | CTTTGTTTCAG |
| 20651 | GGAGACTTAA | AGCATATGTA | TTTTAAAGTT | ACTAGCAGTT | ATAGGATTGA |
| 20701 | CTAGTAGAGC | TATGGCTTAC | ATTAGGGAGG | CAATGCCAAG | GAAAGAATAG |
| 20751 | AATGTGTGGT | TGGCCACCAA | CATNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20801 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 20851 | NNNNNNNNNN | NNCTAACTTT | GCAAGATTCT | TACTGGTTAT | AAGGAGGTGG |
| 20901 | CTTCCACAGT | TCTTGACAGA | TGTTTGCTAT | TATTAGTGTG | TTCTCAATAT |
| 20951 | ATACAATACA | AATGGGACAC | TCTGAGCCGA | GTTTTACTAG | GATCCATCCT |

FIGURE 3G

| | | | | | |
|-------|------------|-------------|-------------|------------|------------|
| 21001 | AGGACAGATG | GTACAGCACA | TTTTCAAATG | GTTTTGTTTA | CCATAATCAT |
| 21051 | CTTTTGGAAA | TAATCAATTG | GAAGTAGAAC | AATTCTTCAG | TTATTCATTT |
| 21101 | CTAAAATCAA | AACCTTCATG | GGATAGTTTT | CTGTGCTTCT | GAACCGTTTT |
| 21151 | TCCCATTCT | CTAATAATAT | CCATCCATAA | ACCTCCGAGT | GGAACAGTTT |
| 21201 | TGGTTGTCTA | ACAGTGTGTTG | TGCTTGGGAA | ATGAATTGAA | GGGAGATGGA |
| 21251 | ATTGAATTGT | TGAATGTGAG | CATTTAGAGT | TCCTGTAATT | TATCCGTNNN |
| 21301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 21351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNTTG |
| 21401 | TAAGATGATT | GGCTCTTATT | CCTTCTATGA | TTCTTAAATA | GGAAAAAGTT |
| 21451 | AGGGATAAAG | ATTTGAGGGA | GTTTCAGGGAA | AGAATCATTT | CCTGCTGAAG |
| 21501 | AGGATCAAGG | AAGACTTCAT | GAACCTCAATG | GTATTCAGTT | GATACTATAT |
| 21551 | ATATATATAT | TATCAGTATA | TCTAGTATAT | AGTATCAGTA | TTGAAAGCCC |
| 21601 | ACAGAGGAAG | GCAATACTGG | GTAAGAGGCA | TGTTTTAGTA | TAGGATTCTG |
| 21651 | TGAGAGCATA | TCTTAAGGGG | CAGTAGCCCA | GACCAGCAGT | GATGGAGTGG |
| 21701 | TGATGGCCCC | GCTGAATCCC | GAGGGATAAA | TCAGTCAAGT | GAAGAGTATC |
| 21751 | CTGGGCACAC | AGAAAAATGC | ATTTAAAGCT | CAGTAGCCAG | AGGGAGGAGC |
| 21801 | ACATTCCGAG | AACTGTTAAG | AATCTCAACC | CAGCATATGT | GGGGTATATG |
| 21851 | TTTGAAGGGC | ATCTGATGAG | AGACGCAGTG | GGAGAAGTGG | GAAAGGACTC |
| 21901 | CTGGAAGGCT | TTTGATATCA | TATTAAATAC | ATTTGCCTTA | ATCACAAAGT |
| 21951 | TGTAGAGGAA | CACTGTAAAG | ATTTTAAGCA | GAGAAATTAA | TATTTTAGAG |
| 22001 | GATAGCTGGT | GGTAATGGAT | CAGATCTAGC | ACTCAGGTGA | ATTTAGTTTT |
| 22051 | GCTTGGAGTG | TTTGTAGTTG | TCAGTGTCTC | ACAGGAGACA | CGGATATAAG |
| 22101 | GGTTGGTTCA | AAAATCGGTA | GGTCTGCATT | GCTGCGAGGA | GTCTATCAGC |
| 22151 | TGGGGCCTCT | TAAGTAGCTG | GCCCATTG | AAGAAGTATG | TGCCTCATTT |
| 22201 | CCCTGTGGAG | CTCACTTCCT | TTAGTGCAAA | TGCCTGGGAC | CTGTAGGAAG |
| 22251 | TCAACCTGGT | GAGAAATTAG | GCTTCATTTT | AAGTAGTGGT | GGTGGGAACA |
| 22301 | AAGAGAAGGA | GAGGGATTGA | AAGATACTGA | AAAGAAAGGA | TGGAGAGGCA |
| 22351 | CACAGTGTTA | GGGCAGTGGC | GGGGGGGCAC | ATGGGCTTGG | CAGACAGACT |
| 22401 | TGNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 22451 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 22501 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 22551 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNCAGTATT | TGTAATGCTG |
| 22601 | GTTGTGAAAG | GTGAAGAAGA | TGTAATCAAG | ATTACTCCCA | AGTCTCTAGG |
| 22651 | ACATCATTTG | GTGAGCTAGA | GAACACAGGA | GAAGGTGCAG | ATAATGATTT |
| 22701 | CAATTTTGGA | TATGTTTCAGT | TTGAAGCGCT | TGTGAGACTT | TGAAGTAGAG |
| 22751 | AAACTTGATA | TGTGCTTGGA | TTGCAGGGTG | AGATCTGAAC | AGAACACAGG |
| 22801 | CTATGATTTT | ATTTCTTTCC | ACTATTCTAC | TCAAACCTTC | CTTCCTTTTC |
| 22851 | CTTCGTCAAC | TTCACTCAAC | ATTGATTATT | GCCACAATA | TCTCCACGGA |
| 22901 | GGAACAGCTT | TAAACAGTAG | AAAAAGAATG | TGGGATAGGT | AAAAATCTGT |
| 22951 | GTTCCACATC | TGCTAAATGT | TATCTAATAG | TTGTAAGTTC | TTGAAAAAAA |
| 23001 | ATCTTTCCAG | TTTACCGTTT | TGGCAAAGCA | TTGTAAGCAA | CTAATAATTC |
| 23051 | AACATTTGTT | TATATCAAGG | GCTGAAAAAA | AGTCATCCTC | TACAAACATA |
| 23101 | TTTTCTTTTT | TTCCCTCTTT | TGTTTGCTTT | TTTTTGTTTT | ACATCACCAC |
| 23151 | CACTCCTAAC | TCTAAGAAAA | TATTTGACTA | AAAGTGAATC | ATTGTTAGTA |
| 23201 | GTGAATTCGT | ACNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 23251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 23301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 23351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 23401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 23451 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 23501 | NNNNNNNNNN | NNNAGCACAT | TTATTTTTTA | AGTTTATATT | CTGTCCTAAT |
| 23551 | ACAATGTATT | AAAAATGAGC | CCTGCCTCAA | GTCTAGCCT | AAATGCCACT |
| 23601 | TTCTCCAGTC | TGCTTTCCTG | ATTTTCACAG | CAGGAAGTCT | TCTGTACTTC |
| 23651 | CTCTACATAT | AACACTTGCA | TGAAGCATAT | CTCACATTGC | CCTACATTGT |
| 23701 | GGTTATTTAT | GGATCTATAT | TATTTGTTCT | GCTAGACTAT | TAGTTTCATG |
| 23751 | AAGGTGGAGG | CAGTGTCTTA | TACATCTTTA | CGCCAGTATA | ACAGCTTCTG |
| 23801 | CAATGCCTCG | TCCAGTTCAG | TTCACTACAC | TTTTGAAGGT | GTAAGAGCTA |
| 23851 | GTCATTTCTA | TTTATTAATA | AAAGGGATTT | AAAAAAAGA | TGGACAAGAT |
| 23901 | GTAATCTTAA | TAAACATATT | GTGGTTTTCA | GGAAACAGTT | TAATGAGAAA |
| 23951 | ATGAATATAT | TTGTCTCCCT | ACCTTATGAC | CTTATTACAT | ATCATAGATC |

FIGURE 3H

| | | | | | |
|-------|-------------|------------|------------|-------------|------------|
| 24001 | TTCACCTCTTG | CTTTAGATTC | TAAACAGAAA | TCAATTTATG | TGCTTGAAAT |
| 24051 | CACATAAAGG | TAATCTGATT | GTGTCATCTA | CTGCTTACAG | TCCATTGGCT |
| 24101 | GTTTACAATT | ATGGGTAGAA | TCTTGATTAT | TTGAGTAGAA | GCCTTTTCAG |
| 24151 | GCTGACATTG | CCTTATCAGT | GGAGGACTGT | TTCTTGTTGT | AGGCCCAGTC |
| 24201 | GCAGTGACG | GCCCTTGGGT | TTCCAGACTT | CTGCTGCTTC | ATGCCTGTTG |
| 24251 | GCTTTTCTTA | TGCTGCTCAG | TCTGAGTCAA | AGGCCCTTCA | CCCATTCCCA |
| 24301 | TGGTGAATTT | CTACTTATCC | TTTAAAATTC | AGCTCACATT | TCCTCTTGAA |
| 24351 | AGTTTTTCCT | GGTATCTTTC | CTATTCCTGC | CTCCCTGCAG | AGAAATGCCT |
| 24401 | TCTTCTGTAT | TCTTTTAAAC | CCTTGACAC | CTGTCATGTC | TATCTTTGTC |
| 24451 | AGGAGACTGT | ACACTACAGA | GCAAGTACTG | TGTATTTAAA | GAGCTCAAAG |
| 24501 | TCTAGTGAAG | GAGCTAGATA | AGAAAAACAA | CTCTGAAAAT | ATGGTATAAG |
| 24551 | ATATATATAT | TTAAAAGTAA | ACCTGTATAT | ACATATAATT | TATGATAGAA |
| 24601 | GTATGCACAG | AGGCTTATGA | AAGCACTGAA | GAGGGTATGA | ACTCAGCATG |
| 24651 | AGGCGTGAAG | ATCTTGGGGG | TGATATAATG | GTCAAACCTGA | CAGGTGAAGC |
| 24701 | CTGTCTAGAT | ATTAGTGTAG | CAGACAAGAG | GGGGAATGGC | ATTCCATGTG |
| 24751 | AAAGTGTTGA | GGTACTCGAG | AGTGTATTTT | TAAGAAGTGC | AATTAAGTGC |
| 24801 | TATATTTGTG | AATGTGATTG | GAAATGAGGC | CAATGGGATA | GGCAAGGGTC |
| 24851 | ATATCATGAA | ATATGGTACT | GAATGCCATG | CTGAGTGTTC | GAAATAGTAT |
| 24901 | AATGATCTAC | ATGAACATGA | TGAGGAGCCA | GTGAATGAGT | ATAAGGAGCA |
| 24951 | GAATGGCATG | CTGAAATCAT | GGGTTAGGGA | GGTCCCTCTA | ACAATAGTTC |
| 25001 | AGGCCTGGAT | TGAAGTCCAT | GATGTTGGAC | ACAGAAAGAC | TGTNNNNNNN |
| 25051 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 25101 | NNNNNNNNNA | AGGTTCTTCA | ACAGATAAAT | TTGGAATAAT | TATCATTTTC |
| 25151 | TTTAGAAAAC | AAATAAGCAA | ACTAAAAACA | CCTGAAGGTC | AAAATGACAA |
| 25201 | ATCTCTGACA | TCATATTAGT | CGAGGCCCAG | TAACAAAATT | GNNNNNNNNN |
| 25251 | NNNNNNNNNN | NNNNNTATAC | TTCAACTCAT | GTATCAGTGA | CTCTTTTAAA |
| 25301 | CATAGAGATT | TTGGTTGCTA | AAAAAGGTGA | TGATGGTTAC | ATTAAAATTT |
| 25351 | TCTGGCCTTT | TGTGATGATT | CTTAGCAAAA | CCTGCTCAAA | ATTACAGTAT |
| 25401 | CAGATGTTTT | AAAAAATGAT | AATTTTCAAC | TGAAAGAAAT | ACTGTAATGT |
| 25451 | ATTTTCATCAT | ATTCTATGAA | TTCTTACATT | ATGAAAGATA | TCTTTTATGT |
| 25501 | TTCAGATATC | TATTATATTT | CCTTTTGTAA | AGAATTTTTT | CTAGCACTAC |
| 25551 | ATCTTTTCCA | AACCTTTTTG | ACTTTCCTAA | GCCTGTGTAA | TTAATCACTG |
| 25601 | CCTGTTCTGT | GTCTTAACA | TGCATCTGTT | ATCAGGACAT | CTCTGCTTAC |
| 25651 | TTCTGTTTCT | CCTTACTAGA | CTTACCTGCT | GGAGGATGAG | AACTTGCTCT |
| 25701 | CCTATTCATT | TTTATATCCC | CAATTTTAA | TATAGTGTCT | TAGACACTGG |
| 25751 | GGCGTTTCCT | CCTTGTCATA | CTTATAATCC | TCNNNNNNNN | NNNNNNNNNN |
| 25801 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 25851 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 25901 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 25951 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26001 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNCCTT | TCTCTAAATT |
| 26051 | ACATTGAAAA | GTTAACAAGT | ACTCACAAGT | TAAATAATAG | TTGCTGTGTG |
| 26101 | ATTCATGTTA | TTGTACTCCT | TTTATATAT | TTCTTCCATG | NNNNNNNNNN |
| 26151 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26451 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26501 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26551 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26601 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26651 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26701 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26751 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26801 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26851 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26901 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 26951 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |

FIGURE 3I

| | | | | | |
|-------|------------|-------------|-------------|------------|-------------|
| 27001 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 27051 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 27101 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 27151 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 27201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 27251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 27301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 27351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNTAC | TCTACTNNNN | NNNNNNNNNN |
| 27401 | NNNNNNNNNN | TTCACATTTC | CAAATAAATA | GTAACACTTT | TTAATGTGTT |
| 27451 | CTTAACGTTT | TATTTGGCAT | TAATCTAAAT | TCCCCTCTAG | CACAATGAAA |
| 27501 | ACAGAAAGAG | AAAGTTAAAA | TTTCAAGTAA | ACTGAAAACA | ACAATAGTGC |
| 27551 | TCAGAGGATT | TTTTTTTAAAG | TGAAAAGGGA | TAGTGCTTAA | TTATGACAAA |
| 27601 | TAAAAGCTAA | CTTGAGATGC | ACATATACAG | ATGCAGCCAC | TTATTTTGGC |
| 27651 | GGGGGACACT | TCAGGAATTA | AAATTTAAAT | AGCGAGAGAT | CAAATAGATA |
| 27701 | CTTGGTAAAT | GTGTCTGAAT | TGGATGTTCC | CAGACACAAA | AATAAAATGA |
| 27751 | GTTATTGACA | GCTCTTGGGA | GACAACATTA | TAAAGACTAG | ACACGTTATT |
| 27801 | TATTTTAACT | CTATGTTCTA | AATTACCATT | GAGTAATTGA | CATTTCGTATT |
| 27851 | TGACTATGGT | TTGTGGTTAA | GTTCTTAATT | GCAATAATGT | TAAATAAAAT |
| 27901 | GTGAAGCCCA | AAGCAAACAA | CAACAAAAAT | TATAGCAATA | CTTCAACAGA |
| 27951 | GGTAATAATA | ATATGCTGCA | TCAATGGTTC | AGAATCCAGC | ATCTACATAA |
| 28001 | AACAAGCAAC | AGGGTAATGA | AATTATTTTC | TTTTCAAATA | TTCTGGCAGA |
| 28051 | GCTACTTTAG | TTTTCTTAAG | TTATAGATTG | TGGTCTTAAC | TGCAACTTTT |
| 28101 | CGCTCCTTTT | AAGAAGTATT | TAAGTTATTT | AAATGTTACT | TAAATACTTT |
| 28151 | TATGTTTTTA | ATCATTATTA | ACATTCTCTA | CCCCTCTCAT | CCTTTCCTGC |
| 28201 | TTAGTATTTT | GTTATAATCT | CACTCTCCCC | ACTTCCAAAT | GCACTCAAAA |
| 28251 | ATGCTGGACT | TTCTGGTCTT | TTTCTGACCA | CCAGAAGAAG | TAGTGGATTG |
| 28301 | AGCATGGAAC | TGGGATAAGA | CATTTTTCCN | NNNNNNNNNN | NNNNNNNNNN |
| 28351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 28401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 28451 | NNNNNNNNNN | TGAGTTCTGT | ATTCATGAGG | GCATTTAATG | TGTCTTATAG |
| 28501 | TGATAACAAT | CAGAGGCATT | ATTATTTCAGT | TTTTACTGCA | GCAGGAGTTT |
| 28551 | AGATTTGACC | ATCAGATAAT | TGACACTATA | GTACAGGGTA | TTAGGAAAGC |
| 28601 | TGGAAAGATG | CTTTCCCCTG | GAGATCTTGG | ACAGTAGGAT | TCTTCTAGGT |
| 28651 | TTTGTTCTGC | TCATGATATG | CAATGGCTCA | GGGCACTTTT | CAGAGGACCA |
| 28701 | CTGTACAATC | TCCTCCCAT | CCACCCAGGT | CCTATTAATA | TAGTAGTGAG |
| 28751 | TTGTTTCATA | TTGCCTTCAG | TGACTTTCCC | TTTTCCAAGG | CCAATTACCA |
| 28801 | TTTGAAAAAG | TCATTCCCTG | TCATATTTTC | TTGTTCTGCT | CAAATGAATT |
| 28851 | TGTCCATATC | TGAATTCTGT | GAAGCTTGTT | GGGTGTAAAA | ACAGCTTTTC |
| 28901 | AAAGCTTCC | AGTTTCATAT | CTATTTGTTC | TTCTGCCCTA | CCCTTGCATG |
| 28951 | TTTCTCTGTC | CCCTTTTCTT | CTGTGAGCAT | AACCTGGAGG | ACAAGGTTTT |
| 29001 | TTGTTCTGTC | TTTAGTGGTG | CCATGTCTGC | ATAATAACAT | GATGTAGATT |
| 29051 | GAAAAAATTA | CAAATGATTC | TTGGAATTCT | AAAGATAATC | TTATTTCTAT |
| 29101 | TGAGAAAATC | CTTCTCAAGT | TACTAACTAC | CATTGAGATT | GGATTGGCTT |
| 29151 | TTGCCTTATT | ACTTTTCAGA | GCCTCCATCT | GCTTGGGTAT | CTCAACATAT |
| 29201 | CCTTAGTTTT | CAAATGTTGG | CACTTTACTC | CCAGCATGAT | TACCTTACAT |
| 29251 | AAGAAACATT | ATAGGACTGA | TGTGGGAGTT | TACTTTTCTC | ATAACTTATT |
| 29301 | TGATAATTCA | CTGCTTATGT | TAGAGTTAGA | AACTATTGTC | CAACTCTCAG |
| 29351 | AGACCCAGTT | ACATCACTTA | AGATGGATAN | NNNNNNNNNN | NNNNNNNNNN |
| 29401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 29451 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 29501 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 29551 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNGGAT |
| 29601 | GGATTTTTTA | GACTAAAAAT | TTAGATTGCA | ATCAAACAGA | AGAGGCTTGG |
| 29651 | ACATGTGACA | TTAAAGAGCT | ATGTTTGTTT | TGCCAGTTGA | AACTTGTTTC |
| 29701 | TTTTTCTAGT | TAAAACAATG | ATTTAGAGGT | TATTTTGAGG | GCACTTTGAA |
| 29751 | GATTATGCTA | CAGGAATGCT | AGAGAGCAA | TTTTGAGAGT | GATTGCCCAT |
| 29801 | TTGGACTTAA | TTCTGGCAAC | TGATTTTGGG | GTAAAATGTC | TATCATTCGT |
| 29851 | TGATTATCTG | TGATTTTAC | CTGGACTTTA | CTTAAGCTTT | ATTAAGTTGC |
| 29901 | TAAACCATAT | TTGGATGCTA | GTGATAGCAG | ACCATCAAAT | ACGGCCCAAA |
| 29951 | CTTCTTGTTT | TGATCCACAC | GAAGGTCAGA | GAAGCAATGC | TGCCTTTTCT |

FIGURE 3J

| | | | | | |
|-------|-------------|------------|------------|------------|------------|
| 30001 | GATAGCCAGT | AGCACCAGCA | GGCACGTTGT | TTACTCCAAA | CAAGAATTTT |
| 30051 | AATATTTTGG | AAGACCACTG | AAAATGGATC | ATTTATACTT | TTTTATTTTT |
| 30101 | TTGATAAGGA | AGGGATGACC | TACTATTAC | AGAGTAATGC | AGTTTGCTGA |
| 30151 | AAAGGTTGGT | TTTTGCTGAC | CTCTGAGAGC | TCACATTACA | GTGGAGTGTG |
| 30201 | TTATTGGAAG | GTGAGCCAGT | TTATACAGAG | GTGGGAATT | ATTTTTCTCT |
| 30251 | AGTTTTGAAA | TGGTTATGCA | CTTTCTAATC | TAGTTATTTA | GATGTAGAAA |
| 30301 | TAAAGTTTTT | TTTTTACATT | CCCCTGAGTA | TATGGTATTG | TAGGTATAAA |
| 30351 | ATAAATTTGA | TGAGTTTTAT | TTCTGTACTA | GGATCTACTC | ATAACTCTTC |
| 30401 | TATCCTAAAT | TGTATCAAA | AGGAAACCTT | TGCTGTCTTG | ATAATAAACA |
| 30451 | CAGACTAAGT | CCAGATTCCA | CAGTCTAGTG | ATCAAGGAAA | TTCAGGAATT |
| 30501 | GTATTTAGCT | ACAAGTAACG | TGACACAAAG | AACAGTGCCC | TAAATTGCTA |
| 30551 | GGATGGTGAT | TTAGGGTTAA | TATTATGACT | TCTTGCTCTA | CCCTCTTTCC |
| 30601 | ATTCTCAAGA | TCGCCTCATG | GTCTATAAGA | GGCCACTGTG | GTTGAGTCAT |
| 30651 | GACATGCACG | TTCTAGGCGA | GAACACAGAA | CAAGCCGTGC | TCTTCAGCCC |
| 30701 | CCTTCTTACA | CAGCATTCCA | AAGCCCCACC | CCATCACTTC | TGCTTTCATC |
| 30751 | CCATTTAGCC | AGACTTAGT | CATTGGCTG | CCCATATCTG | CTAAGGAGAA |
| 30801 | TAGGGAACAT | AGTTTTCAT | TACTAACCCA | TCCCCACTC | NNNNNNNNNN |
| 30851 | NNNNNNNNNN | NNCCATTCTA | TATTAAAGCA | GAAGGGAAAG | AGATATTGGT |
| 30901 | AAGAATCCAG | CTGGCCTTTT | GTGATCTGTG | TCAGCCTTTC | TTTTGATCTC |
| 30951 | ATCTGCTGTT | TAAAGCACTT | TACACTGTAG | CCCCACGAGA | ACACTTTGCA |
| 31001 | CTACTAAGA | AGCAGTCCCC | TTTGCTGCGC | CCCCCCCCCA | CCACTTTGCT |
| 31051 | TATATTCTGA | AAGTCTTTTG | TTTCCTATTG | CACTGCTCTT | ACCTCTAACA |
| 31101 | CACCTGCCTCT | AACACACCAA | CCTGCAGTTG | TAGTTATTAC | ACACCCCTCT |
| 31151 | TGGTCTTTTC | ATCTCTCTAT | CACAGCCCTT | GTTGTGGTTT | AGCCAATATA |
| 31201 | TTTTAGTTCC | ACAGCTAAAT | TTTCATACCC | TCTATGACTC | TCTAATCCCC |
| 31251 | TGCCACACTT | GCCTACTATA | ATACATTATA | TATATAACAA | ATGTTTGATA |
| 31301 | CGTATTTATT | GAATTCCATT | CCAGAACTAA | TGCCAGCAAG | ATAACTTTGT |
| 31351 | GCTATATAGG | AGAATATCTT | TTTGTGCAAC | AGTTTCCAAA | GGGTTTCTTT |
| 31401 | TTTCTAAGAA | GAAAGAAATT | GATTGTATCA | ACTTTATGAG | TATCCTACCG |
| 31451 | CATTTAATAG | CCATTGGCTA | ATCTAAGGGT | TCCTGGTTAC | TTCATGAAT |
| 31501 | AGCCTATCAG | ATGGAAGTGC | AAACAACAGT | TTGTTTTGAA | ATAGGACTCC |
| 31551 | CTAAACATGG | AAGAAACATT | AACAGTGTTG | GCCTGTTGGA | ATGTGTGCAT |
| 31601 | TTGATGTGCT | CAAGATTAGG | GCACTCTGCT | TGAGAACAAA | TAACAAAAAA |
| 31651 | GGGAGAGGAA | ACAATAAAAA | CTTTGGTCCT | ATAAAGCACC | TGAAAGTACT |
| 31701 | ATAAATTGAT | GGTTCTCAAG | CTGGTCAGGG | GGTCCAAAGG | CTACAGCCTG |
| 31751 | GGGGCCTCAA | GTTTAATTGT | TTTATAAAGT | GTCTTAAATA | AAATTTTTAT |
| 31801 | GTTTTTAAGT | GGTATTTTTA | AAACTACTTT | TTCTATGCTT | GAGAGAGTTT |
| 31851 | TCCAAATTCC | AATTATTTTA | AGGGGTTATT | TCCTGGACTT | GCACTTAACG |
| 31901 | ATTTTGAGAT | GTTTACATTT | TTTTCAATAT | GGCATTCTGT | GTGCCTCAGT |
| 31951 | GATACATGGT | TATCCAGGTT | GCATGCATAT | ATAAATGTTA | AGATTTATGG |
| 32001 | AAGGTCATCT | TTTTAGATTA | AAAAGAATTT | TTTTAAGCTG | GTATTTCTTG |
| 32051 | GTGATAGGGC | CTAGAAATTA | TGTAGAGTGG | CTTACTTCTG | GAAGTTATTT |
| 32101 | TAAACTGCAT | ATAAACCATC | CGCCTAGTGT | ACAGTTGGCT | AAAGAGTAAT |
| 32151 | ATTAGAAGGC | CCTCCTGGAC | AGTTTATTTT | ATTTTATGGA | TATGAACACA |
| 32201 | ATTGTTTTCC | TTTGAATTTA | ATGCCATGTT | TAAAATCAGA | TTTTAAGAAT |
| 32251 | TTTCCAAGGG | CATTTCCCTA | TCATTTACAC | TCTGCTTGTT | TTTTCTTTCT |
| 32301 | GTAGTCTTTT | ACATTAAATA | CCTCCTACAG | AGCACTGCCT | AAGGATTTGT |
| 32351 | GGTGGTACAG | GGTCCAGTTG | GGATGACAAA | CAGGCAAGGA | AGGCCTGGAA |
| 32401 | GTAAAAATTAG | CAAAGAGGCC | CTGTGGAATG | GAAGGTGAGG | GAAGGGTTAG |
| 32451 | TGACAGTTGG | GGGAGGAAAG | GTAGAAAAAA | AAATAACATG | CACATCAGTT |
| 32501 | TCGCAGGAGT | ATTAGAGTCT | TAAAGGAAAC | AATGTTTGAT | AATTATCAGA |
| 32551 | GAGGAAACTG | GGAGCATAGA | GCATATCCCA | GAATGGAGAA | CAGCATGGGT |
| 32601 | TAAAATGGGT | AGGAACAGGT | GTCAGGAGCT | TCAAANNNNN | NNNNNNNNNN |
| 32651 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 32701 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 32751 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 32801 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 32851 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 32901 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 32951 | NNNNNNNNNN | NTGATTTTTA | TCAGGCAGTT | ATGATGATAA | ATGTATGGAA |

FIGURE 3K

| | | | | | |
|-------|-------------|------------|------------|-------------|-------------|
| 33001 | ACTTCCCATT | TCCTAGAGCT | AAAGTGCATG | TTTCTCATTC | TGAAATGTAG |
| 33051 | GGAAACATAAT | CATCTGATAC | CACTCACCTG | ATTGTTTCTC | ACTCTTCCTC |
| 33101 | CACCATTTTAC | CCATCTCTTT | AGCTTAATGA | GTCCCCTGTG | TATCTCCCAA |
| 33151 | CTAAACAGCG | GCTTACTTGC | CTGTGAAATA | TTCTTCTCTT | GGGTAGTCTG |
| 33201 | CTCCCTTCTC | TGTCTACTCA | TGCTTCAAGA | TTCAACATAA | GCCTCCTCTA |
| 33251 | TGAGGCTTTC | TGCACGTATG | TATATGGATT | TGCTTGTGTA | ATGATTTCTT |
| 33301 | CACAGATTTT | ATATTGCTGA | TAAATAAATA | TTGTTTTGAA | TAAGAAACGT |
| 33351 | GGTTTTGTAT | TTTTATCTCG | ATTGTAGACT | CCTTGAGACC | AGTACCATGC |
| 33401 | TATACAATTA | TTTTTCATCT | ATTATAGTGT | CTGGCATAGG | GACATGCACA |
| 33451 | TATTTGGTAC | AGAANNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 33501 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 33551 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 33601 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 33651 | NNNNNNNNNN | NNNNNNNNNN | NNNNNTAGAG | AATAAACGGA | ATAAATTCCA |
| 33701 | ATTAAACAGT | GAAAATATTT | GAGTGATGAT | TTACAGAATT | TTAAAGTCTT |
| 33751 | GAGAAAGTGG | AACTCAGTTG | ATGGAAAAGA | GTGAATGTCA | AACTGAGAAC |
| 33801 | GTCTATTTTG | CTATGTTAGG | GCATAGAAAG | GCCATTATAG | ATTGAAAAGC |
| 33851 | AGTTGTAATG | AATCTTAGGG | TTAAGAAAGA | TCAAGATACA | GAAATATCCT |
| 33901 | TGAATTGAAG | TTCCAAAACA | ATGTTGTTTT | GGTTTTTGTT | TCATTTTGAA |
| 33951 | TCCTTTCATA | CTTAGGAATA | CCATTTCTAG | TAAAATAAAT | ATTTTATGTT |
| 34001 | TAGTTAGAAA | TTTATCTGTA | TTTCATACAT | ACTTAGTACT | TTTGGACTAA |
| 34051 | GAACGTGCTAT | TGAAGTATA | TTTGAGGGAA | TGAGTTTGAA | ATTTTGGGTG |
| 34101 | CAGGACATTA | TAAAGTTGTA | ACTATGAATA | AATTTAGTTC | ATGCTTATGC |
| 34151 | ATAGTTTAC | CTAGTTTAT | TTGTCTATTT | GAGTATTGTC | CTTGAATTTA |
| 34201 | AAATTTTTTT | CAGCCCCAAC | TGATACACAC | ACATATACAT | ACATAATACA |
| 34251 | TGTGTGTGTG | TGTAGCTTAC | AGAGTGTTTA | TAGGAAACTG | ATTTTGTATA |
| 34301 | CTTTGGCTAC | TTTGTTGTAA | GTTCTAGTTT | TTTTTCTTTT | ATTATTAAAC |
| 34351 | TAGTGCACGA | CATCAATGCT | ATATGATTGG | TGTTTCGTTG | ACCTAGAAAT |
| 34401 | AATGCATGCC | ATCTTCTTTT | CACAGCTGTG | TGCCAACCCAC | GATGCAAACA |
| 34451 | TGGTGAATGT | ATCGGGCCAA | ACAAGTGCAA | GTGTCATCCT | GGTTATGCTG |
| 34501 | GAAAAACCTG | TAATCAAGGT | AGGAAAACAG | TCTGACATAA | ATACACAATC |
| 34551 | GAAGACACCT | CTATCACTCC | CAAATTAAAA | ATATTCTTAT | CTCAAACCTAC |
| 34601 | TTTCCATGGC | TATTTTTCCA | AAATATGTGA | GCTGCTATTT | TGCTGNNNNN |
| 34651 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNTTTTA | TCAAAGGACA |
| 34701 | ATTAAACAAA | TTGTATCCTC | TTATTCTCTG | ATACTAATTA | AAATGTATTT |
| 34751 | TGAAGAAAAG | AATCCCTGCA | TCAGTAATTT | AGAAGCCTTC | TGGTACTCCT |
| 34801 | TGTTTTTCCA | CTAACTAGTT | GGAAATCCTT | GGCCAAATAA | TTAACCACCT |
| 34851 | CAGACCCCAG | GTAAGCTTTC | CCTTTAATGC | CAAAGTCAAG | TAGGGGATTT |
| 34901 | GATTTGAAAT | TTGGAGTTTC | CTCCTAACTC | TGAGCCCTTC | GATTCGTGAT |
| 34951 | TAAATCTCCC | TTCAACTACT | GACCAGTTTG | GAATGTTTCC | ATGATAAATA |
| 35001 | AAAATGATTA | ATTTAGCAAG | CACTTTTTAA | AAAAATCAGC | ATCAGTTGTT |
| 35051 | TAAAGCAAT | ATTATTCAA | CACTCATAGT | GCCCCAGGTA | TTTTATATAT |
| 35101 | ATAATCCATT | TTTATCTTTA | CAACAACATT | AGGAAGTAGA | GCATTATTTA |
| 35151 | TTATAATNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35301 | NNNNNNNNNN | NNNNNTTTT | GGATAAAAAA | ATATATTTGG | AAATGTGATC |
| 35351 | CTCGAACTCA | TGCTACAAAG | TCAGACAAGG | CTGTCTTGT | TAATTAAATT |
| 35401 | CAGTTAAAAA | TTTCCATTAT | TNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35451 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35501 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35551 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35601 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35651 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 35701 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NCATTATTAG | GAATTAAACT |
| 35751 | AAAGAAAGAT | TAGGGTTAGA | TTTCACTAAA | TAGATAAGTG | TCAAAAATAA |
| 35801 | AAAAGATAAG | TCAATTTTTA | CTTATTTTTT | AAATTATACT | CCCTATCATC |
| 35851 | TTAAATGTCA | GGTGAAATAA | TCATGGTGTC | TAATATCCC | TTACATACTT |
| 35901 | ACTTGACCTC | ATCTGATATA | GAAATGATAT | TGCTGAAATA | CTACTGTTCT |
| 35951 | TCAGTGCTAG | ATACTTATTC | CAAGATACTC | CTTGAGGTAT | GTCATGTAAA |

FIGURE 3L

| | | | | | |
|-------|------------|------------|------------|------------|-------------|
| 36001 | CTAATGATTT | ATAGACACAG | TTTTTTTTTC | ACTATTTATN | NNNNNNNNNN |
| 36051 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 36101 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 36151 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 36201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 36251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 36301 | NNNNNNNNNN | NNNNGACACC | GTTTTTACAA | TCATCCAGAA | CATTGATTTT |
| 36351 | TGAAAACATA | CTCAAAATGG | TGGTTCATGA | ACCATCAGAT | GAAATTATTC |
| 36401 | ATGAACCATC | AGATGAAATC | CACACTAGCA | AAAATGACAT | GTCCCAATGG |
| 36451 | TAGAAGCTGC | TATTCTAAAG | TGTGTTCTTG | TATTATCTGC | TATTTGATTT |
| 36501 | AGAACTGATT | TTCCCAATGC | TTGTTTTCTC | TCTAATACTC | TTTTTAACTT |
| 36551 | GAAATTTACC | AAACATACCT | GCATAATCTT | TTTAAATAAG | TGATGCTCTT |
| 36601 | ATTATCTCAG | TTTGCTCCTT | AAAAAACTCC | ACTTGATTTT | TCTTCCCCAG |
| 36651 | CATAAGTTTG | CAGGTAGCAG | TGTTCTGGTT | ATTGGATGCC | AATGTTTCATC |
| 36701 | TTAGAATCTC | AGTATACTTT | TTTTAAAGTG | GTGAACATAA | GCCCTCAGAG |
| 36751 | GTTTCATTAA | TATATCAGCA | CGGTAAAATA | TTGTTGTCCA | AATGTGAGGA |
| 36801 | TATAAAATAT | AAAGAAGACG | ATTAAAATAA | GTCACTTTTG | AGATTGCAGA |
| 36851 | GAAATTCATT | TAAATTTCTT | TGCAGAAGCT | GCATCAACTT | TAACATGCTT |
| 36901 | TAAAGACATG | CTAAAAGTAT | TATTGAAACA | AAAGTTCATC | ACTGGGTAGC |
| 36951 | TCCTATTTTC | AGAAAACCAG | TTTAACATGT | ACTTTTTTTT | CCTGTAACAC |
| 37001 | ATACCTATTT | CTCTAAAGAA | AAAATCGAAT | GTATCAAGTT | AAGATCTTGC |
| 37051 | TCCCCAAAAC | CATATTCCTT | AAAGAGAAGA | GGATTTGATT | AGTAGCAAAA |
| 37101 | TGGGATTTTA | CACCTGCAAA | AATAGTGCTG | ACACTGAAAA | TGTAATTTCA |
| 37151 | AGTCCAGTTT | TGAAAATAAA | ACTACAGCTT | AAAATAAACT | TAGGTGTTTT |
| 37201 | AATCATCTCT | TATTTTGCCC | TCCTCACAAA | AAAGCAGTGT | GGCAGGTTCC |
| 37251 | TGATGCAAG | GTTNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 37301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 37351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NAGTTTCATC | TAATGGTTCT |
| 37401 | TAGTCCTGAA | CACGTGGAA | GCTTGAAAGG | GTCCACCAG | TTCTTGGTAC |
| 37451 | AGCATCTTAG | AGAGTCTCTG | CATCCATAGA | CTTGAGGGCC | CTGTCAGGAT |
| 37501 | CAGGGAAGCT | GCCATGTGTG | GCCTGGTTGA | GTGTGAGAGC | TGCCTAGAGA |
| 37551 | CTTCCATAAA | AGTTGTTAGA | GAAAATTGTT | GATAGTGCCT | ACATAGCAGA |
| 37601 | TTAACTTAAA | CTGTTTTCAT | TCAGGCCAC | ACATTTTAAT | AAAGTAGAAA |
| 37651 | ATATGCTTCA | CAGATAAGGG | AAATCAAACA | GGCTCCTTTT | TTCTGGAGGA |
| 37701 | GAGAAATGTC | AAAAAGAATT | AAATTTGAAA | TAACTTTACA | GAAGTGGAAA |
| 37751 | TTAGCTTTTG | ATTAAAAGTA | GCTTTTGGTA | TATGACAGGT | ATTCAGTGAG |
| 37801 | AATTTTGTAG | CGAGTTATAT | ACTTTAAGAA | ATAACCCCCA | GAAACTTGCA |
| 37851 | TCATGGTGTA | AACAGCTTGA | ATAACAAGT | GCTTAACCAG | TGCCTTTAGA |
| 37901 | GCTGCCTGGG | AAACAGCCAG | AATACCAGGG | CAAGCTGCAT | TTTGGAAGCTG |
| 37951 | GTTTAATTTA | GTAGCCTTGC | CACAGGCTTA | GTGTGATCTG | CTTTTGGTGG |
| 38001 | CTTGATCTTC | CCCCTAAGT | CATTTTCTGG | ATTTGTTACA | CCTAGAACTG |
| 38051 | TTAGGAAATT | ACAGCTTGG | GCTGATCATT | AACATACGT | ACTCTACAAG |
| 38101 | GCACACGTTA | CCTTTCAAAG | CAGATGAAAT | TCTAACCTGA | ATTCTGGCAA |
| 38151 | GATTCTTTGA | TCATTTGCTT | CCTTTACTTT | TACTTTTATT | TATGCATATT |
| 38201 | TCCCTCCTCC | TTGAGTTTCT | GTACCAACAC | AAACCTCTTT | TTCCCCCAGC |
| 38251 | CAGNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NTTTTCCATT |
| 38301 | GGGGAATGAT | TTTTAGTATG | TAAATATATC | ACTGCATATT | CTCAGAAATG |
| 38351 | AAAGACATTC | TTAGGAATTT | ACAGTGACT | TTATAATAAT | TTCAGAGAA |
| 38401 | AATATTTATA | AATGTCAATT | CCTAATGTTT | TAGCATGGTT | TATGTTTCAT |
| 38451 | ATGTTGAATT | CTTTATCATA | AGGAAAGAAT | TGGAGTCTTT | TAGGTCAGAA |
| 38501 | CCAGATACTA | ATTTTGTGGA | CTAGTTACAT | CTGAAAGTTG | ACTGCTTTGC |
| 38551 | TAAGCACAAA | AATCTAAGGG | CTTAACTCT | AATATTAAGG | TGGTTACCTA |
| 38601 | CAGCCGTAGG | TTTTGAAAGA | TGTATGGTTC | CNNNNNNNNN | NNNNNNNNNN |
| 38651 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 38701 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 38751 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 38801 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 38851 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 38901 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNGAAG | TTGTATGCTT |
| 38951 | TCAACTCCCA | AAACCAATGT | TCATTTGATT | TGGACATTTA | ACTGGCTATT |

FIGURE 3M

| | | | | | |
|-------|-------------|-------------|-------------|-------------|------------|
| 39001 | AATAACATAA | TGGTGTGAA | ACAGAACGTG | ATTGTTAAGT | CTCAACTCTC |
| 39051 | TTCTTTGAGT | CTCACCCATC | TTATATTTTG | GTAAGCTAAT | AATGGCAACA |
| 39101 | GTTAACATTT | TTTTGCATAA | AGCTTCTCAT | GTATGTTCCA | GAAGCATAAT |
| 39151 | TTCTTACATT | TAAATGATAA | TATTGACGAT | CTATTATGCT | GTTATTTAAA |
| 39201 | TAAAATCTTT | ATGTAAATCA | TACATTTGTG | CATAATTTTG | GAGTATTTTA |
| 39251 | AACTACTATT | GTACAACGTG | ATAGTGTAGG | GAAAAAATC | TTAAGTTGGT |
| 39301 | CGTTTACAAT | TCAGTGTCC | TAAATAGAGC | AGGTGGATTA | TAATCAATAT |
| 39351 | TTGTATAAAT | TGTTTTTGT | TTTCTTCCTT | CTGTGTATGA | CTAAATAATA |
| 39401 | CTTTTAAAAA | TGATGATCTC | TTGTTTGAA | TTTTTTAAAA | AATTGTATGA |
| 39451 | CGATGGATTG | TTCATTTATA | ATGTGTGTGT | TGTCCTTGGG | AAACAAGATG |
| 39501 | ACAGTGATAC | ATTTTAAAGA | AAATTAAAGG | AAAAGAGAAA | CCAGAAGAAC |
| 39551 | CAGAACTAT | TGTTTATAAA | GTATCTAATG | CATTCTTTTT | AAACTCTACC |
| 39601 | AACTGTATGT | AATTGTTTAA | TATATACTTG | CAAAGATTTT | TAGGGCTAAA |
| 39651 | ATTGACATCA | GTTCAAAGTT | GACTCTTATT | ACCTTCTTCC | TGGTGTGAAG |
| 39701 | ACGAGCACAT | CCCAGCTCCT | CTTGACCAAG | GCAGTGAACA | GCCTCTTTTC |
| 39751 | CAACCCCTGG | ATCACCAAGC | CACAAGTTTG | CCTTCAAGGG | GTGAGCGTGC |
| 39801 | ATTGTCCCAG | GTGGTCTGCT | CTTCCCACCT | GCATGGCTTC | CTGCAGTTAT |
| 39851 | GTCACTTCTT | GCAAATGTTT | TGAAGACTAG | CCCAGGAGTG | TCCAGCCTTC |
| 39901 | AGCCTGAGC | ATGAGCAAGC | AGCAGGAATT | GCCACCTGGT | GCTGATTCTC |
| 39951 | TGTGCTCTGA | AAAGCATGGA | GCCCTACTGT | CCTTTTAAATG | CATCGAGTAT |
| 40001 | CATTGCTTTT | CTTCTCTTTT | TCTGAATTTT | GAAATATGTT | CTAAAAACAG |
| 40051 | ATAGCAAAAC | AACCTTTTTT | ATATGTAGGC | AAATTCAGGA | GTTCCAAAGC |
| 40101 | AACTGAAATT | GGGGGTTTGA | TGCCTTTTGC | CAAGGAGATT | TTAGAAATCA |
| 40151 | AGTCCTCTTC | CCCCTCTTTC | AGGTATAGAG | TTTTAGGTTA | TCTGAAAATC |
| 40201 | AAAGCATACA | TTAGTGACCT | TCTTTTGGTA | TACATTGTGT | AACATCAGAT |
| 40251 | AGCCTAAGTT | CTGGTTCCAA | GAGAGCACTT | AGTTGCTATT | TGATTATAAG |
| 40301 | CAAGTCATTA | ACCAATCAGT | AACAGGGAGA | CAGTAATTCC | TGCCCTATTT |
| 40351 | GCCTGTGGGG | TTGGCGTGAG | AATAGAAACA | CCAAAAATAT | TAATGTCTTC |
| 40401 | GTCATGCTGG | AAAGCTTTGT | AGGAAAATAA | TGTGGTAGCA | TAATTTTACA |
| 40451 | GTTCTCTTTT | TAGGTCATTT | TATCTCTAAC | CCATTCATAG | GTTTAAAGAC |
| 40501 | TTAGAATAAC | TGAATTAATA | TAGTTGAAAT | TATTAATCAT | TGTCTCTGCT |
| 40551 | GAGAAGAAAT | TTGATGCATT | TGTATATTTT | CTACACAAGG | AATTAGGGCA |
| 40601 | AAAGAATTAA | TTTCGGTCCN | NNNNNNNNNN | NNNNNNNNNN | NNNNNGTAGT |
| 40651 | TGACATATTG | ACATCTGTTT | GCCTATGACA | TTGCCCGATT | TAAGCACCAA |
| 40701 | AGCGGGGAGA | AAGTCCAAAT | GTGTTAAAC | AAATGAAAGT | TTATTTTAA |
| 40751 | AAAAATCCAGT | AAGTTAGTTA | TCTCCCGATT | TTTCAAGCTA | CTTTTCAGTG |
| 40801 | TCTGCATGCT | AATAAAATTT | CTGATTTTTT | TCCTGAAGTT | TAATAATAAT |
| 40851 | GCTGTGCAGC | TTGCATTTCC | ATTCTGGACC | AGTTCCTTTC | TCCCTCTCCT |
| 40901 | TCTCCCATCC | AAGTCTCCTGA | ATCCACCAAC | CTACTGAAAT | GTATTCCTGA |
| 40951 | CAATAGTAGA | TGCACTTTAA | GACTTGTATA | CATAATAACT | GAAGCATTTG |
| 41001 | AATGTAAGTG | GTTTATTGTA | AAGTTCTATC | CTTTAGTGTA | AGATAAGCAC |
| 41051 | ATGAAATAAT | TCTTAAAT | TTTTATTTT | CTATTTTATA | GATTTTCCTT |
| 41101 | ATATATATTA | TAAACCTCCA | GAGAAAAAGG | AAGATAAGTA | AATTTTAAAT |
| 41151 | AAAAACACCA | AAGTTTTATT | TCTAGGTTCT | TTTATCAACT | TTTAAGATTT |
| 41201 | ATTTGAGACA | GTATGATCAA | TGACTTCATT | TTGTTCTGCT | TATTATTGTA |
| 41251 | GGAGTATTTA | CTATAATTTG | GAAGTAATTT | ATTTTGAAT | TTATTGCTTA |
| 41301 | ATTGAATGAT | CTCCAATAGA | TTGTGATAAT | GAACACAGCA | TTTATAGAAA |
| 41351 | GCAGCACATA | TTAACTTACT | TAATATGGCA | CTAGGTCAAT | GAGAAAAGAA |
| 41401 | GGTAACATAA | TTGAAGACAA | GAAACTCTTA | AGAAAACTGA | GGACAAAAAG |
| 41451 | GCTTCTCACC | AGGACACCAG | ATGCATTTAA | TCTTTTGAAG | CTCTGTACTT |
| 41501 | TAGGAAAAGT | CTGATATTTG | GCAAATTTTG | ATAAACATGG | ATGACTATGG |
| 41551 | AATCCTATTT | TATAGTATCT | GAAGTGGCTT | TCATAAGGGT | CATTGTGAAG |
| 41601 | TTTTTAGGAG | ACACCTGCCT | GTGGCAGATG | GGACAATGAT | GGCAGTCACT |
| 41651 | AGTGATATTA | ACACCAGTCA | GCTGTCAGGG | AATATCATCC | AGACCATCAG |
| 41701 | CAGCTGGTAG | AGTACAGCTT | TCTCAATTGC | TTTCCATGTT | TTGGATACTT |
| 41751 | ATATGCCCGT | TAATAACAGG | TAAAAATAGCC | AGTACATCAT | TTCCACATTT |
| 41801 | ACCCATTGAA | TGTTGCATGT | TTTCTTCCTT | TCACATATTC | ATACAGTCCA |
| 41851 | GATTTTTTTT | TGGACTCATG | ACAGCACATT | GGCTTTTCTT | TCCTTTCAGT |
| 41901 | TTCATGATTC | TTAACCCCAA | AGTGCTTTTG | CCATGGGAAC | GGAAGGATAA |
| 41951 | ATTTCCGGTCG | AAGCATATTCT | CTAAAACCAC | CAGCAGCTCC | ACTCCAGATT |

FIGURE 3N

42001 ATGAATACTA CCATAGCAAT CCTAAGGGGC CCTTTATTGG CCCAGACTGG
42051 AGCCAAATTG TAGAAGAGCT GCCCAACTG GGCTTCAGGT TTAGTTCAGG
42101 GACATAGTGT CTGAAGAAAT TTCCATCAGC ATAAATACTC CTTGTTTATG
42151 AGCTGCCTGA AACTGTAAAC ACCGAATCCA TTCCCATCAG GAACTCACAA
42201 AAGTTTCTGT TATGCTTTGA AAATAAACCT AGGGATACTT AACTGACATG
42251 TAAAGAAAAA TCCATCAGTA TCCTTTTCTC GAGATATAGG TTTTGATTTT
42301 ACTGTGTTAT GTTGTGGTTT TGTCTGTTT TTGTTTAGTG CAGAAAAATA
42351 TTAAAAACAC AAAGCCTTTC TCTAACACCA TATTTGGTTT TAAATGCTAT
42401 TTGCTATAAT ATCAAGAAGA TTTATCAAAG ACAGGTGTGA CTCTGCAGGA
42451 CCATATTAGG AACAGTCATT AACTCCTCTT AGAAGATGAG AAAAGTTTCC
42501 CTCTCATTTT TAGTAAATTC TGAAATACAA AGTAGAAGAA TATGGGAAGG
42551 GTAGAAAATT TCTCAGTCAT TCTTCCTTGA TTTTGTGCTC AATTAACAGA
42601 GGGAGAAAAC ATTTTTGATA CTGTAATCTG CATGGCGCTT TCTGTGAGGA
42651 AATTTTTTTT GAAATATTTA AAATGTGATT TTGTTTTAAA AACCCCTAAC
42701 TAGTCAAGAG ACAATAAAGT AGAAAATTGA AGAACGCATA GGAGATGACA
42751 AAAGTATTTG GTTTATTTAT TTAACCTTTC CGGGATTGTC CCAGCTCGGG
42801 GTCTGCTTAC AGTTGCATTC AGGGCAGATA AAAGACCTAC TTGGAAAATC
42851 AGTAAGATAT GAAAATTTTT TTAGATGTAA TAATTATTTG GTGGTTTTAA
42901 AACATAATGG AACTTGATGA TTTAGTTAAA TGAAATATAC AGATTTCTAA
42951 TAAATGAGCA GATCAGATTG AAATAGATTA GATTTGAAAA CAATTTTTGT
43001 CACTAGACAT ATTATCTATA TTTTATTTC AATGACATGG ATTAATAGAT
43051 AATTAAATTA TGCTAAATGG GTACTTATGC TGGTGGTTCC AGACAAACAT
43101 GTGGTCACAT TTCTTTTTAA TGAAGCATAT ATTTGGGCA TTACTATTTT
43151 AATGTTTATT TGTCATTATG AGTTTATGT TCACTTGAGG CAGTGTTTAA
43201 TTTAGGTAA AACCTTTTCA TAATGTAAAA TTTGTTGATT CATTTTCATAT
43251 TGATACTTAA ACCAATAAAC CTGAAATACA TGACAGTCTA TTACATTTTG
43301 TCCATTAGTG TTTCCATATG CCTTTTTTAA CTGCGGAATA AATGAAAATA
43351 TGTGTTAGG TTTTTTAAAG TTCATTTTGT GACTTGAATG CATTACTAAA
43401 GTAGGAAACT GAAGTTTCTT CTTTAATCTG AAATCATATA AAATCTACCA
43451 GATTGAACAA GAAAAACTAC AATATTGATC ACTCTTAATT TATTTCTGTT
43501 TTATCCTCTG AAAATATAGC CTATGTAGCC TTCCTTTGGG ATGGGAAGGT
43551 CAAAGCAGTC AAATGTTTTA AAATCTGTAT ACTTCCTGTA AAATAAAATA
43601 TTCTCAAAGT CCAAGCCTTA AGGAGGCCAA TGCCTTATCC ATTTAAAGTA
43651 AAATATTCTC ATGTAATATG TTCTTTAAAC AGAGAAAGGA AGACATTGAA
43701 CGAATAGGGT TACATTGTCA GGATTGGATC TAGCAATAGA ACCCAATAT
43751 TTTGAGAATA TTGGCAAACA GTTGTAGCT GATAGAGCAT CTCTGCCAT
43801 GAGGACGTTA GTATGCTGCT GTTCTAGGG TTGATAAGGT AGGCACCTGT
43851 TGTTGAGAAT TTTCTAACT GGCCTTTATA AACAGAGATC CATAAAGGTC
43901 ATGGCGACTC CTGTTGCCC CTAGTGTAAC GTACCTTCAA ATTAATATTG
43951 CATGAACAGG ATCTATCTCT TATGTAAATG ACAGTTACTG TAGCTTGACC
44001 TATTTTTCTC TTTCATTTTT TAATATAGAA AAGAATTTTA GATGGCTTAC
44051 ATTAGTGCTG ATGTTTGTGA ATTTTCTGCT TCAACCCAC ATTTTCTCA
44101 GATCTAAATG AGTGTGGCCT GAAGCCCCGG CCCTGTAAGC ACAGGTGCAT
44151 GAACACTTAC GGCAGCTACA AGTGCTACTG TCTCAACGGA TATATGCTCA
44201 TGCCGGATGG TTCTGCTCA AGTATGTCAA GAATCTTAAC TGTTTTATAA
44251 GTGCTTTGGG CTGTTTCTG TTGTGCTCTG AGAGCTTGCT TTTGTGAAAA
44301 TGGCCTCCGG GGTCTCCTA AACAAGATGT GTGCGTGTGT TTATACTTTT
44351 GCCTGAGGAA TTGAAAATCA AAATAAGAGG CACCATTTTC GATGTATAAT
44401 ACTCTTCTAT CACGGGTGCC AATATTAAAT TGATTAGGAA ATGGTGTTTA
44451 GAATAAGAAT TACAAAATA AGTATGTTTT CTAATTTTTT TTTTAAACAT
44501 GTTTACCCAT GTCTTCTATC TGGGCATAGA AGAATGATTC CAGGCTAAAA
44551 AAAAAAAAAA AAGAACAATA CTGACTATTT AGGGACTGCT CAGACAGAAT
44601 CCATCAGGTC TGTTTTGTAA TTTTAAATCA TTCATAAATA TTTTCTTCAT
44651 TTTTATATTT CATTATAAAA GCCTTTAGGC TTTTATAGAA TTTTAGACTC
44701 TAATAATAGT CTAAGACTTC TAAAAACAA TTCAAATAA AAAATAAATG
44751 AATATGATAC CAGAAAAGTA GGATTCATGT TTATAAGGAA GAAATGCATG
44801 TTTTACCAT CATTTGGATG TAAAAAATGG ACTTTGCCCT AAAATTCTCT
44851 ATAGGCTGAC TCCTGAATTG TGGTATTCAT ATGATATCTG ACAATTAATG
44901 ACTTGATTTT TATTTTTACT GAATTATCAT TTAAGTGAAT TGCATAAAAT
44951 AAAGTTAGCT AATTTTATT AACATTGCAT AAGTATAGCA TGGAACTTTT

FIGURE 30

45001 TGGAAATTAGG TGAATACATG TTTAACATTG TGCAACTCAA TGGGAAATCT
45051 GCTGTTCCTCC TAGAGAAATT TCATGGGCAT TTGAGACAGT TACTTGGATG
45101 ATTAGTTAAG AATAGCGTTT AGCTTGAGTA ATCTGGAAAA ATACCTCAAT
45151 TCTTCACTTT CCTTGGCCAC TGGAAAAATT TCCAGATAAT CTATATATGA
45201 TAGTTTTATA TTTGTCTTTT AATGCTTTCT TTCCTTCCCC CTTTTATTTT
45251 CTTTAGATGT TTGATAGTTA AGGCCTCTTT GTCTTCATTG TTGTTCACTG
45301 TCACTGATTT CAGGACAAGT GAAAAAACAT AGCTATTTCC ATTACAGATT
45351 TTGTTTTCTA CCTGTCTGAG TCAGCCAGTC ACTCTCTTTC CAGGTGCCCT
45401 GACCTGCTCC ATGGCAAAC TGCAGTATGG CTGTGATGTT GTTAAAGGAC
45451 AAATACGGTG CCAGTGCCCA TCCCCTGGCC TGCAGCTGGC TCCTGATGGG
45501 AGGACCTGTG TAGGTGAGTT GTAAAATCAA GCATCTCTGT CAGCAGCCTC
45551 TGTAGGATAA AGGGAGAAAG TGAAAGGTGA TGGGAATAAG GAAAAAAAAG
45601 GCAATTACTT ACATCAGATA ATTAGCTATC GTTCAGAAGA TATCAGATGT
45651 CTCAGAAGAG GACATCTCTG TGAATGGATA ATGGGAGCGT TGTTTTTTAA
45701 AAAATGAGAA ATANNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
45751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
45801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
45851 NNNNNNNNAT TGTATGTGTT GTTCTTCAGG CCTCAGATTT TTTTCTGGTA
45901 AATTAGGAAG GCAGTGTAGG ATTGAGAGTC TGGANNNNNN NNNNNNNNNN
45951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNA AGTGTGTTGGT GATTAGAGTN
46151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGTCTATT
46351 AGCCTTAAGG TAAGACCTGT ATCATTAGCC ATGTGAAGAA ACAGCCAAGC
46401 AGCTGAAGAC ATTCAGGGC ACTGCGGTGA GAGAGAGGTT GGCCTGTGTG
46451 AGGTTTTTCA AGGAAGAAGG TCCATGTAGC TCAAGTACAG GGTGTGAGTA
46501 TGAGGGGTGG TGGCTGGAGA TATAATGGGT GAAGTAGGTG AGGGTGATTT
46551 GTTGCAGGGC CTTCAGGAC ACAGGGGAGG AATGTGGATT TAATTCAAAG
46601 TGCAAGTAGA AATAGTTAAA AGATTTTGAG CACAGTAATG ATCTGATTGT
46651 ATGTTTTAAT GGAATCATTT TGACAGTTGG GTTGATATG CTTTGGAGGG
46701 GGTAAGTAAC AGCAAAAGGA ACTGTTTTGA GGCTTTTGTA ACAGTCTAAC
46751 CCAAGAGATG GTTGCTTCAA ATAGGGTAGT GGTGGTGGAT TAAGAGAAAA
46801 GTAGACAGAT TCAAATACTT TTTTTTAGAG GTAGAATTAA CTGGACTTGA
46851 TGATATGGGA GTTAAAGGAA AGAGGGTTGT TTCTAATTCA AGAATGACTC
46901 ATGTTTCTAG TTTAAGCAAC TTGGTAGATA CAGGTGCAGT ATGTGCTAAT
46951 ATATAGAGTG CAAGAAAATA ATAACCTGGC TTAGAAGGCT GGGTTGAAAT
47001 TTGGATATGT GGAATTTATG TTGCCTGTAG GGACAACCAG CTAGATACCG
47051 TTTATGATGA GCCACTATAT GCTAGGCAAC CATTTGTAGT CAGTTATTTA
47101 TTTTGAAACC CACTCTGTCT TAACATAACC TTGCTCTTCC TAAAATGCTA
47151 TTGACTTATG TTTCTAGATG TTGATGAATG TGCTACAGGA AGAGCCTCCT
47201 GCCCTAGATT TAGGCAATGT GTCACACTT TTGGGAGCTA CATCTGCAAG
47251 TGTCAATAAG GCTTCGATCT CATGTATATT GGAGGCAAAT ATCAATGTCA
47301 TGGTAATGAA ACCCAACCAT TGCTTTGTGT TGTTCCTTCC TAGAGCACTG
47351 AAAGGTCTCG TAATTGTGGT GATGGCTGGA ATGTCAGGGG CAGGGGAGAG
47401 TACTGGCGTT AAGTTAAACC AACAGACATC CAGTTTAACC ACTGGTAGTT
47451 CTCAGTCTAC ATGTAGTTTA TTTCTTCTGT TTATCTGCCA ATTTTATGTA
47501 GATCATCACA TTGCCAAAAA AAATCATTTT TGAAACTGTA TATATTTTTT
47551 ATGTCATCAT ATTTATCTCC TAAATAAGTC TCTTCTTTTC CTAATTTCTG
47601 ATGCAGACAT AGACGAATGC TCACTTGGTC AGTATCAGTG CAGCAGCTTT
47651 GCTCGATGTT ATAACGTACG TGGGTCCTAC AAGTGCAAAT GTAAAGAAGG
47701 ATACCAGGGT GATGGACTGA CTGTGTGTG TGAGTAGCAC TTGTCTCTCA
47751 GCTTTAAATT CTAGCAGGAA ATACAGGATT ACACAAAGGC CATTGCTAGG
47801 GAAAATAAGG AATAAGATTA TCAAAGAAGT ATAATTGTCA TAATTGGTTA
47851 TATTTGTCTT TGATTTCCAC AAACAATAAA ATCACTTGCT CAGGTACTTG
47901 TAAAAACTAA GGACTCAGTA ATACACTATA ATCTTAAGAG TATTTTAATC
47951 TCTTCACTGA AATCTCTCAA TATTTTCTTT TTAGCTAAAA AGAAATTATT

FIGURE 3P

51001 TTATATGTGC ATAAC TATTT TCTGAAATTT ATGTTTTTGT AATTTTCTCC
51051 TCACCCCTC TCTCCATCTC AGATACTGTC TTTTCTCCC TATTACTCTC
51101 CTTTTTAAAT TCTCATCATT GTGATCATAA AATGGGATTT TTAATTTATG
51151 AGTGCTTAAG TAATTATGGT ATTTACATTA TTTTGCTGCC TTAGGATAAC
51201 AGGGAAATTT GGCTATTTAA TGTAAGATGA TACGCTAAAT ATTTTTTTCA
51251 TTATGATGAA AGAATACATC TTTCTGAGAA TTTTAAAAAA TCCTTTCTCT
51301 TTTTAAATG TTTCTCTTTA TTTCCAGTAT TCCTCTCTAT GCCAATACAT
51351 ATATGAATTA AAAATGACAT TGAAGTTCAC CAATAAAATT TAGTGCAATA
51401 AAATTGGGGA AATACAGAGT TCCAATGATG TTTGGGAGCA TTCATTATAA
51451 GAGAGTGTGA TCTTAAAGAC ATGCTCTGGG AAAGCATTGC CTTGATCAAA
51501 TGCCAAGGCT GTTGCATGCC ACAGATAGCA TTGCTGCCTT TAGAAACTCC
51551 TGCCAAAAT TAAATTCAC TCTCATTTCA TCTTCAGCGA TAGCTGCTTA
51601 TTAAGTCTAG CATGTGTGAG AATGCTTTAG ATGCTTTTTG ACTTGCTGTC
51651 TTGTGGTTAT AGCATATATT CTAAATAAGG CAAAGGCTCT AAGTTTTTAA
51701 TTCAGGACAA ATAACAGTG TCTATTCAT GTGAAAACAG TAAGGAACCT
51751 TAGCACTATT TCAGTTCAC ACCATTTCTT TACTGAATAT TTTCTTCACT
51801 TCTGGCATAA CTCTTAAAT CATCCTCATT CTCCGAGACA GGTCTGAGAA
51851 TAGAATTTAG CTCACCTCAT TTCATAAATT TTAGGTGAT GTTAATTATG
51901 GCCACTATTT GGCAGTCTCC TTCAAATAGC CCTCCACTTT GTGTTTTTCT
51951 TTACATTGAA AGATATTTTT AAATTGANNN NNNNNNNNNN NNNNNNNNNN
52001 NNNNNNNNNN NNNNNNNNNG AAAGATTTCT TTTAATGGAA GAGGTGTGTA
52051 AAGAGAGCTT TTCTTATAA TGTATGTGCG TTGTTGCCCT GAGAAAAGAC
52101 TGTAAAGATA TTCTAAAAGA AAATCAAGGA AGAAAAATAT TATAACAAGA
52151 ACACATCTTC AGCCCAGACC TCTCCCCCAA ACTCTAGACC TGGATGTCAG
52201 CTGACTGCTT AACCTATCCA GTTGGATGAA AATAGACGTC TCAAACCTCA
52251 CATGTATAGA ATTTCTCCTT TTCTTCCCTT AAACATCCTC TACTCTCAGC
52301 CTTTCTTATC CCAGGTAATG GGACCTCTAT CACCTGTTGC TAAGGTCTAA
52351 TAATGTGGAT TCATCCTTAA TTACTCTTCA TACATACAGT CAATATGTAA
52401 GGAAATCCTG TTAGCTGTAC CTTCAAATTA TATTTAAGTG TGACCTTTTC
52451 TCACCAACTC CAGTGCTACC ACCCTGGCCC AAGCCATCTC TCCCCTGGAA
52501 TACAGGCCCA AAATCCTTTA TCCAAATAAC TTATGGAATA TAGCATTTTT
52551 TAGATTTTAG AAAATCAGTA AGGTACACAT ACTATATAGT ACACACTGAA
52601 GTAGTGAAAC ATGCTGATTT TCCTCTAGTG CTTTACTGT GAACGTATCA
52651 ATGTTAAGGA AAGGAAAATG ATATTAATC AGAGATGATG TCTCACAGCA
52701 TATATTTACT AGCTTGACA AATTTTTTAA TGTTAGCAAG ATTTAAGGCA
52751 AATTTTTTAC TTATATTTTA ATTGATCCT ATGATGATTA TTAAGAAAAA
52801 AAGTAGTTAT CTCTAGAAAG TATAAATTGA GCTCTTGGCA AATGTGAGCA
52851 AAACCAGAAA TCAGATTTTT TTAAAGTTAC GTGTACATTT GTCTATAGAG
52901 TTATAATTAA AAGTATTGTG CTCATAGCCA GTCTGTGATC TGGGGCATAT
52951 TATTATATCT TCCTATGCCT TAAAACAATG TTTCTGTAA GTATAAACAA
53001 AGTACCATTG AACTGCAGT TTTTGTCCAT TTGCTACTTC CAAAAGAAAG
53051 GGGTACAAA CAGAAAAGTT ATTTAACTT TAAGCAGTTG GTTAGATATG
53101 GTAGTTATAA TGAGCTTTGG TATTTTAATA TGGGGCTTAA AAATTTTCAT
53151 CTAACCAAC ATTAATTATC TAAGTGTGAT ATCCTTAGGA GGTCTTTGTT
53201 AAGTCTCCC AATTATATCC AGCATCTCCA AAGGTGACTG AAGTCCCTTC
53251 TAATAACTCA TCNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
53951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN

FIGURE 3R

54001
 54051
 54101
 54151
 54201
 54251
 54301
 54351
 54401
 54451
 54501
 54551
 54601
 54651
 54701
 54751
 54801
 54851
 54901
 54951
 55001
 55051
 55101
 55151
 55201
 55251
 55301
 55351
 55401
 55451
 55501
 55551
 55601
 55651
 55701
 55751
 55801
 55851
 55901
 55951
 56001
 56051
 56101
 56151
 56201
 56251
 56301
 56351
 56401
 56451
 56501
 56551
 56601
 56651
 56701
 56751
 56801
 56851
 56901
 56951

| | | | | | |
|-------|-------------|------------|------------|------------|------------|
| 54001 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 54051 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 54101 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 54151 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 54201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 54251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 54301 | NNNNNNNNNN | NNNNNNNNNN | NNTCATCCCA | CCACTGGAGC | ATTTTTCCTT |
| 54351 | ATATTTAACC | TAAATTTATG | AAACTACAGT | TTATGCATTT | TTAATATTGT |
| 54401 | CAGATACTCA | GGGTTAATGG | ACACAGTTGT | TCATTGTCTC | CAGATCCCAT |
| 54451 | ACTCCATGTA | CTCAAAGACA | CTTTCAGGAG | ATGACTCACT | CTTCCCTGGG |
| 54501 | CCCNNNNNNN | NNNNNNNNNN | NNNNNTTTTC | TTAGAAATGC | TTTTTTTCAA |
| 54551 | CCCACTTAGC | AATCTTGTTG | ACTGGTCCTT | TGAACATTCT | TTATTTTCCA |
| 54601 | AATCTGTCAC | AAATGATCGA | GTCAAAGATA | TATAATAGTC | AAATCTGACC |
| 54651 | CAATACTGAT | CACGGGGGGG | TTTCTTTTAA | CATGTTTGT | AATCTAAATC |
| 54701 | ATTTCTCTTT | GTTGCAAAGT | AGAGATTAT | GGGCAGCTAT | TTAATCTCCA |
| 54751 | GATGTTTTTT | CCTCAAANN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 54801 | NNNNNNNNNN | NNNTGGAATT | TTATATTGGA | TTTAATTTTT | TTCATTCTAA |
| 54851 | GTAATCCTTC | ATATGCTTAA | CTAGCTAAAT | TTTACTTGTT | GGTTCTGATA |
| 54901 | TTCTTCCCC | AAATAATCAA | AGTTAATTCA | GATTATGGTT | TTTTTTCTTA |
| 54951 | TCCTGATACA | GCTTAATATA | GCCATTGAAA | ATATTCTCTA | GAAAATTCTA |
| 55001 | AGTCTAATAT | ACTTCTAAAT | TTTAAGCATG | ACACTGAAAA | ATAGTGAAAA |
| 55051 | GCATAGAAAA | GGAAATAGAA | AATAGCATTG | AAACATTTGG | AATTAAGGGG |
| 55101 | TTTTAACCAC | TACTAAAGAT | AGAATATACT | GAACCAGAAA | ACCTAGTCTT |
| 55151 | GATTTCTACC | TCTATTTATG | AGGTACATGA | CTTTGGACTA | GTATGAGTNN |
| 55201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 55251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNTCCC | AGCTATCCTT | TCTCACTGAG |
| 55301 | TAGCTTCAAA | GATCAGAGAT | AAGGGATCTA | GTGTTTGTGA | ATTCTGAGTT |
| 55351 | ACTATGAAAT | GTAAGAGATT | AATGTGTTTT | GTTTATCACT | GTAATAAACC |
| 55401 | TTGATGGGGA | GAACTTATGT | GATTTAAGGT | TGCAACTTTT | TTTGTGTGAT |
| 55451 | GTGAAATATT | AAACATATTT | TTAAATAAAN | NNNNNNNNNN | NNNNNNNNNN |
| 55501 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 55551 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 55601 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 55651 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 55701 | NNNNNNNNNN | NNNNNNNNNN | AAAAACAATT | TTAGCTTGAG | GGACCCTAGG |
| 55751 | TTTTTTTAGG | TTATATGACT | AGAAATATAA | TAAGATGTAT | TTACAGGGCT |
| 55801 | AATCCGGTGG | GCATGATTTG | AGCAACACTA | AAGTTATCTG | ATGTAGCGTA |
| 55851 | AAATCTTGT | TGTTTCATTA | ATGAGACAGA | CTTGTGCTTC | TCAAAGTGTG |
| 55901 | TTTTATAGAG | CATTATTTCT | CTGAAATATT | ACACAGTTGT | ATGTGTTAAA |
| 55951 | AAGTTCTTAG | GTCAAATACA | TTAAGAAGT | CCTAGGTTTA | GCTAAGTTAA |
| 56001 | ATAGGTTTCT | TAACTGTAGG | ACTTCTCAAT | ACTCCAATGT | GCTAACAATA |
| 56051 | TTGCAAAATCT | CTAAGAGAGG | CATATAATGT | ATCTTTTCCC | AAGTTGTTTT |
| 56101 | AACCATGGGA | TTATTTGAAT | CTTGGGATTA | GAGTTTACTG | GAATCAACTT |
| 56151 | TAGGTCAGGC | TGGGCTAAAC | TAGGAAGTAT | GGAGACCATA | CTCCACAAAA |
| 56201 | CCCACTATCC | CTCACTAGTT | AGCTCTCGTT | GCCTGGTTTT | TGAGACTAAT |
| 56251 | GGGTTGATCT | TAGATTTAGG | AATGAGGGTT | GGGGAGGGCA | GAAAGGTGCT |
| 56301 | GCACAAACTT | AGCCAGAAAA | GCAGCCATAA | GTGGTGGTAT | GCCTGTATGG |
| 56351 | ACACCAATAG | TTTGCTGCAG | GGGATTGTTA | AAAAGCCACT | TTTCAAAGAC |
| 56401 | AAAACCTCTG | TGTCACAGGG | GTAGGAGCAA | TCCCAGGGCA | AAGAACTCA |
| 56451 | GAAAAACTTT | CTGCAGAACC | CAAAGTAAAG | GAAGTTAGGG | GGACATGCTT |
| 56501 | TCTGTTCTAG | CAATATATAA | TGAAATCTTA | TATAACTATT | AAACATGTTT |
| 56551 | TATGAAAAAT | ATTTAATAAT | TTTGAAAAAT | ACCCATGTTA | CAAAAATGTT |
| 56601 | ACCTGAAAAG | TAGAAAGCTG | TACCTGCCAT | TTGACTCCAA | TTAGATATAA |
| 56651 | AATACATGTC | TGCATATGTC | TGGATGTGTG | TAAAAGTCTG | AAAGAAAAAA |
| 56701 | GACCAAAATA | GTAACCATGA | GTATATCCTG | TTTATGATAT | TTTGGGCGAT |
| 56751 | TTTCATTGTC | TTTTTGTGTT | CTTTATCCAG | TAATTTTTTA | ACTGTTCAAC |
| 56801 | TGTAATATCG | TTGATAACAG | TAATTCATTA | TATCATTAAT | TCCTTCAGAG |
| 56851 | AAGGGGCAGA | GTAACAGAAC | CCCACTCTTC | TCAGCAGGAA | AGTGAGGATT |
| 56901 | CTGATTCTTA | ATCATGTTTA | TTTTATTTTG | GACATACTGA | GTTTTAAAAA |
| 56951 | TTCATTTGAA | GTTTATAGAA | CTTCTATTTT | TAACCTTTCA | GATTTTCCTT |

FIGURE 3S

57001 TACAATTGAA TGTGATAGAT ATGTCCAATT TTAGCTAGGA AATTTGTATT
57051 ACGAGAGGGA ATATAACATC ACTTGGACCA GGAAACAGAT TTTTTTTCCC
57101 CAGTGTGTGC AGAAGCCCAT TTAACTAAA ATAACATAGG ATGTTTATTA
57151 ACTAAAATAA CATAGGATGA TTATCTTAAG ATGTAGATTC CTAATCCTTA
57201 TGACACATTT ACTAAATCAG ACTTACTAGG GAAAGTATCT CTCAGTCTTT
57251 ATCNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57451 NNNNNNTCTG AGTAGGCACA CTAGTCCCC AACACTCAAC CTTTTATTCA
57501 ACCAGCGGTG AGCTGGATGT GAACATGACA GACCCAGTAG GGTTCCAATG
57551 CCTGACAACC TGCACCTGTC AGGAAGAGCC CCCTCCTTTC CTGCTCCCCT
57601 GCAACACATG GTTANNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN ATAGCAGGGC TCTTCTAAA
57901 GCATAATGGC AGCTCCAACC CAGCAGAACC ATATTCTGGA TGGAGTCAGA
57951 ACTAGCCAAA TGTAGGCTAT TTAATAATA GCAATTGTGC CTAGGCATGT
58001 GTCCACCTGA CTCATGTAAA ATAAAAAGAA AAGGCGTTAA TAGAAGCCAG
58051 ATTATAGAAC ATAATGTTAT TAATATGTAT CAGGTGCTCT TATGAGTGGC
58101 CATATCAGAG AACCAGCCTC ATTGTTGCTG TTATTACGCA ACATGAGACT
58151 GTGCTGCAAT TCCAAGTACC AAGCAGTAAG AAGAAAATTG TTTTCTTTTA
58201 CTGATTGCTG CATGTTGTCC TGAGGTTTTT CCCCTCATCT CTCTTCTACA
58251 GATTTTACAG AGTCTGTGGC TAGTTCAGTC ACATTGTTCT TAGTCATGGA
58301 AATATTATGG TCTTCTGAT TCTGTGTAGT GATAAGTAAA AAGATTGTTT
58351 CTGCTGAGGG GTGAAAAGGT CTTCAAAGTA GTTTGCTTTC TTGTAAACAG
58401 TAAGACCTGC AGAGACCCTG AGAAGATGCC TGATATCTCC TTGAAAATTA
58451 AATTTCTGCT AGTGTTTTGA AGGAGCGAAT TGTCACCTCT CACAGGTTAG
58501 GATCTGCTGC TGTGAATTCT GAAAGTTTTT AAGATTTTTG ATTTATATTT
58551 TAATTAGATC CTATGGTAGC TAGTAAAGAA AATATCATTC TCTCCAAAAC
58601 GTATAAATGG GCCCTTGGTG ATGTGACCA AAGTCAGAAG TCAGATTTTT
58651 TTTTGAAGTT ACATGCACAT TTGTCAATAG AGTTATAATT TACAAGTATT
58701 GTCCTCANNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNCTAC
58901 TATTGATGTT GAATTCATTT ATATAGTAAT GATAACATTT CCTACTTAAT
58951 TCATAAAAAG ACAGCCTATG CTGTTTTCTT GTTCTGAGTT TATATGTTTC
59001 TCATGCTTTT TATTATGGTT CATTACAATT TTAATGTTAT TTTTAACTAA
59051 CTAGATCCTT TTGAAACAAA TTGGTTTGCA AGTGTGAGCT GTTAGGTGCA
59101 CAGAGAAAAA TGAAAATAGA AACTTGCGAT TTTATTCTAG GCTTGTTACC
59151 AAATATTTAG AATACTGTGT TTTATTTAGG TGTTTATAGT CTCATTAGAC
59201 AGTTGTGATT TTAATAATAGA GACCACATCA TCTCAACTTC TTTACTGTGA
59251 AAATAATGAC AATAGTCTTT TCAGAGATGA ATCTGTCTAG ATGGGAAATT
59301 TACATGATTG ATCTGATGAG NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59351 NNNNNNNNNN NNNNNNNNNN NNNNNNNAAA ATAATCTGAC AAGTAGTTTC
59401 CCCAGAAAAT CTGATTTAGT AAATGTACCA AAAGGATTTA GAAATCTACA
59451 TCATAAATAA ACATTCTATG TTATTTTAGT TCAGACCCTA TTTTAATTCA
59501 GACTTCCTAT GGGATAAAAA CTTCCATTCT TTCTTTAAAT AGATTCTTTT
59551 GGCTTGAGTG CATTTACACC TGTCCCAAC AGCTGGTGGG CTCTGCTCA
59601 CCCTAGACGG TGTTCTATGCT GCACTCAGTC CAAGCAGCCC TTATCAGAGA
59651 GTCTTCTTAC CACTTGCATT CTGGTGCGAA GGACTCATTC CTGGCAGAGC
59701 CTAACCTTCA TGGGAACCAT TGGCTTAGAA GAGAGGAGAA GNNNNNNNNN
59751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNTATCC

FIGURE 3T

60001
60051
60101
60151
60201
60251
60301
60351
60401
60451
60501
60551
60601
60651
60701
60751
60801
60851
60901
60951
61001
61051
61101
61151
61201
61251
61301
61351
61401
61451
61501
61551
61601
61651
61701
61751
61801
61851
61901
61951
62001
62051
62101
62151
62201
62251
62301
62351
62401
62451
62501
62551
62601
62651
62701
62751
62801
62851
62901
62951

| | | | | |
|-------------|-------------|-------------|-------------|-------------|
| TGTCTCCGGG | CAGTGTCTCT | GAGTCTTTCT | AAAACCTTTAT | TAAAAGTTCC |
| TACAAATACAT | AAAAGAGAAA | TAGGTATGTC | ACAAAAATTG | GGCTCATTGT |
| TGATTAGCAA | TGTCTATGCT | CCACCTTTCA | CCAGAGATTT | AACTTTTTTG |
| CCAATTTTGC | TC'TTTACTGT | CAGCCCAAAG | GTGTCTGTGG | AACTTTGTAG |
| ATTCTTCCTA | TGGGATGAAA | GCTTTGACAA | ACAAGGTCTT | ATTCTGTGA |
| GTCCAACAAC | CCTTTTCTTT | TGGCTTATAA | TGAATAGATG | GCTAAAACCT |
| TTCTCTAACA | TGGTTAAAGA | AAT'TTCAGAA | ACTAATTTCC | AGCCCTTTTA |
| TTTGCTTTTC | TCTCTGATCA | CTTAAAATTG | TGATGCATAT | GTCCTGCATC |
| TGTNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNCATC |
| TCATAAAACT | CTTGTCATTC | ATCCCTACGC | TCCAATCTGT | GTCTCTAAAA |
| GCCTCTAGGT | TTTGCCACCA | GAAACAGCCT | TCAGAATATT | GGAATAATTA |
| CATATGTACC | ATTCTCTTCA | ATGAACTAAA | CTCCTAAGCA | TAGAAGTAAT |
| TTAAGACAGT | GTTATTTTTA | ATCATTTAGT | CCAACAACAA | ATGCCAGAAC |
| TGGCCTGGGT | AATGCCATAG | GAACCACCTT | GTTTCCTATG | AAACAGGAAG |
| AGAATTTGCA | AACCTACCTT | TACTAAATAA | TGCTTACATT | TTGCTATAGT |
| TACTCTTGGC | AAGAAGTTGA | GCAGTGGGAG | GTGTATATGA | TAATATTTAC |
| ATTTACTTCT | CTGTGCTTAC | TGTTAATGTT | TTCTGGGTAA | AAATATGCAA |
| TTGACTATTT | GGGAAACTTC | ATTTGTGAAC | ACAGGTTATA | TATGATCACA |
| TCCCTGGGGA | AAATACATAT | TTAGTAGAAA | GTGCCTGGCC | AGCATTTCTT |
| ACAAGAACTT | TCCTGCTCCT | TTTTGTGGCT | GATTTTACCC | CTGACTCCAG |
| GGCCAGTAGT | CCATTAGAAA | GTACTGTGCT | CCTCAGCCCA | GTCCAACCAT |
| GCCTTTATCC | AGCTGTCACT | TGCTTCAGTA | CCTGTACATT | CCCTTTCTTT |
| TAGTTTATGC | AGAAGGCTGT | AAGAAGCAAC | AGGCAAGACA | TATCTTTTGG |
| TGGGTAAAAC | ATGGACCAAT | GGTATAAAGA | TTCTGGGGAC | ATTTTCTAAA |
| AATATATGAA | CATTCATGGT | TGATGTTAAT | TTTAGTCATA | ATTCCCCTAT |
| GACACCCACT | CTCACCTTCT | TCCTGTTTCT | TCTCCAAGCT | ACTACCAGTA |
| ATGGTCATTT | CAGCATGCAA | ACCTGATCTT | GGCAATCCTC | TGTTAATCCT |
| TCAGAGGCCG | CCAGTTTCCA | CCAGGGATAA | CGTCCAAAAT | ATTTACAAGG |
| CTTTAAAAGC | TCCTGCACAC | TCTGCTTCTA | TCTCCAATTG | CATTGATGCT |
| CCTTCTCCT | GCACCTTCTT | GCTGCCTCCC | TCCCCCTGTA | AGGTCTACAG |
| CAATTCTGAT | GTTTCAGTCC | CCTCTCTTTT | CATTCTCCAG | AGCCATAGCT |
| CATAGTACTT | CTTCTTATGT | CATAGAGAAA | GCATCACTTC | TCAGAGATTT |
| TTATGACGTG | TATCCTCCCC | AGCCTAGCTT | AGCCAGTCCC | TCACTCCCCA |
| CTCCCGTACT | CACTAGCATC | CTCTGCTTCT | TTAAAGCACA | CATTACCTGG |
| GGAAATTGTT | AAAAATCENN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNAATGAGT | GGATGGAGTA |
| GAAGGAGGCA | GGATGCATGG | ACAGATGGTG | GAGGAATTGT | ATTGCTAGCA |
| CTTCTCTGAT | TC'TTTTTTAG | TTTCTAAATT | TATACCATTG | TAATTCTAAG |
| AATTAATGGT | TCAACTGCAG | AGAATAGCAG | GTATGCATTT | GTTTTTATGA |
| AATATATATA | TTGTGTATTT | AAGTTGAATA | TTAGCCCAAA | GTATCAAGCA |
| GGGGAAAGAA | CAATAGAAA | AATCATGTAC | CAATGCTCTAT | TCCTGGGNNN |
| NNNNNNNNNN | NNNNNNNTTT | AATTAACATT | TTCTGCATTG | ATCAAAATAG |
| CTCCCTTGAA | ATCAAAGAAA | AGTGTTTGAA | TTTCACAAAC | ATATTTCAAT |
| TTCATAGCCT | TGCGTTCAAC | AAGTATCAAG | CCTATTCTAA | GTTCTCTTTG |
| AGTAACCAAA | ATACAAATAA | CAGACATATA | ATTGCTATTT | ATATGTGATC |
| TAAACAGAAC | CTCTTTTATT | TAGATGTCTT | AGGTAAGTTT | TTTATTCTAA |
| TTTATATTTT | TTCTTTTAC | AATCCTAAAA | ACATAGAATT | ACAAAAAAG |
| TGAACCAATA | CAAAATGAAA | AAAAAATTTT | ATTCTAATA | CTTAGATTAC |
| CAATATTTAG | CATATTGGCA | TATTGGTGTA | TTTATCCCTC | TTTAGTCCTT |
| TTTTTTTGGT | TTGTGTTTCT | TGTACATGAT | TGTGATAATA | ACATAATACT |
| GCTTGAATGA | TTTGTCTGTC | TACTTATTTT | ACTTAATATT | TTACATAAAA |
| CTCTAAAGTG | TATCGGGGGT | GGGTAGTTTT | TTTTCTCTTG | AAACTTCAGT |
| GGGATGCTCT | TAGTAATCCC | ATACTGGTAT | GTGTGAGGAA | GGAAAATTAG |
| TTAAATAATT | TGGTATGGTT | ATAGAGTAGA | GCACACAAA | TTGTAAGAAC |
| CAATAGCTTC | TGAGCACTTC | TGCTCTCTAA | ATTCTTTGGA | ATGTTTCCCA |
| GTGGATTGTA | ATGAAGGTAT | ACATGATCAT | CTGCTGCTAA | ATTAAATGGT |
| TCTTAGAAAC | CAAGAACCNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| NNNCATCTGA | CATTTTACAA | GTGTGTTACTT | GACTAATTCT | TTGTGTTGCC |
| ATCTTCTAAC | TNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |

FIGURE 3U

| | | | | | |
|-------|------------|------------|------------|------------|------------|
| 63001 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 63051 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 63101 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 63151 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 63201 | NNNNNNNACA | GGCTCCCGGT | GCCAGGAAAC | CTTTACATCA | ACACTCGATT |
| 63251 | TGCCATTTGA | TAGTCCTTCA | TCTGGGAGGA | AAAAAAAAGA | CGGAGGGGAG |
| 63301 | CTTGAAAAAC | TGTCATAATG | TCCCTGGAAT | ATGGTACTTT | TAAGAGTTGA |
| 63351 | GCCTATTCCA | TTTTGGAGAT | GATTTATATA | AGTTACAACA | AAAGAAGGGG |
| 63401 | ACAAAAACAT | GATTGTTCTA | TGGAGTTTTT | ATACTTTCT | GTCACAAGAA |
| 63451 | AGCACGCTTG | TCTACAATTT | TGTAATATTT | CTAGTAAATA | AAAGAGGCAC |
| 63501 | TCCCGCTCTC | AGAGACCAAA | ATAAGGAAAG | TGTAATTGGA | TGTCATTGCT |
| 63551 | GTCAGTCAGC | TGGGCTATAA | AAGAGAGAGT | GGGGTTGCCT | CATCCCCTGG |
| 63601 | GTATCCACAG | TCAGCTGTGT | CCCTAGAGCT | TCTTTTCTTT | CATTGCTGCC |
| 63651 | CAGCTGGGTA | TATTGCAAGT | ATGGATTATA | AGAGGGGAAG | GGACTTCACT |
| 63701 | GTTTTAACGT | TTGAAACAAA | AAGGAAAAAA | CTCAGAAGTA | GTAAGCTAAA |
| 63751 | AACAACTTGT | GCAAACGTTT | TGGGATTATT | ACTTAATTTT | AAGAATTTTT |
| 63801 | GCTAAAAACA | ATAGGAACAT | CGTTGAAAAA | AGAACCCTT | TGAGTGATGA |
| 63851 | CTGTTTCATA | GTGTTCTACA | TCTCCTACTC | CCTGCCTTAT | AAAAATAAAG |
| 63901 | CTAATTAATT | GAAGTTCTGG | CAAAGAGAAG | GGAGTATTCC | TTGGCCTTTG |
| 63951 | ACCANNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 64001 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 64051 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 64101 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 64151 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 64201 | NNNNNNNNNN | NNNNNNNNNN | NNAATGATTT | AAAGAAATCA | ATTATTTTTT |
| 64251 | TAAAGTAGAT | TCTTATGCTT | TGTCCCACTT | GTGTCCTTTG | AGGAAAAGTG |
| 64301 | TGAAATACTA | CCTGGGGTTC | TAACAGGATC | TTTGNNNNNN | NNNNNNNNNN |
| 64351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 64401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNA |
| 64451 | CACATGTGTG | TAAATAGCTG | CAATTCTTTA | TGAAAGTGTG | TAATAGGGGC |
| 64501 | TTATAAGCAT | GAACCTTGGA | GGTACAAGAG | GGAAGAATTT | GATTCTGCCT |
| 64551 | GGGGCAGAAT | CACTCCACAG | AGGCCGTGAT | ATTTATGCTG | GGCCTTGAAG |
| 64601 | GAAGAAGAGA | TTTTTGGCAG | GTTAATTAGC | AGGAAAGAGA | GCATTGGGTC |
| 64651 | TGAGACGTGA | AAATGAATGG | CATGTTTGGA | CAGCATTGTG | TATTTTGGTG |
| 64701 | GGGTAGATTG | TAACTCTGG | CCCAAGAAGT | TTGAGTTTGA | TCTTAAAGGA |
| 64751 | GCCAGGGAAA | CTGTTTAAAG | AAGGGAGGCC | ATGGTAGGAA | ATGTATATTA |
| 64801 | GAGTGCAACA | ATGTGCGGAG | TTCACTGGGG | GTGGGGAGTA | CATGCTGTGA |
| 64851 | TTAGAAACCA | AACCATTGCG | TTCCATCATA | GGTATTTTCA | GTTAAAACAA |
| 64901 | CCTTGATTTA | AAGATCATAA | TAATGTATCT | TTCGTAGAGA | TGAACCTTAG |
| 64951 | AATTTCATAG | AGATGAATAG | TTGGCTAGAA | ATTGGTTGCT | CTATTAAAAT |
| 65001 | GTATTTTCCA | GATAAAGACT | TTTTAAGTTA | TTTTTGTGTA | ATTTTAGAAA |
| 65051 | CTTTTGAAAA | TCATTTTTAA | AAAATGTACC | CAGATCAAGG | TTGCAGTATA |
| 65101 | AATCACTGCA | GAACAAAATT | CGTGAATCTT | CATGTGATAT | ATAATTAAAT |
| 65151 | ATTTTGGATT | CATCCAGATG | AATAAAATGA | CTAGGCCTTT | TCCTCCTGCC |
| 65201 | CTTGTTAAAG | GTAATGCCTC | TTTATCAATA | CTTACACAC | ACACACACAC |
| 65251 | AAAAACTCAA | AACCTTTCTC | TTGCAGTTCC | ACGCAACCT | TCAAATGACT |
| 65301 | TGTTTGAAAT | ATTTGAAATA | GAAAGAGGAG | TCAGTGCAGA | CGATGAAGCA |
| 65351 | AAGGATGATC | CAGGTGACAC | TTACACTCTT | AGTGCCATCA | TAATGACATT |
| 65401 | TTATTGTTTC | TCTCCATGAA | AATAACTTTT | AAATGTAGAT | CGTTTGGACC |
| 65451 | ATTTGGGAAA | ATTACACCTG | TTTTCTTTAA | CAACTCAGAT | TTTCTTTGTA |
| 65501 | GTAAATCAG | TTGAGCNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 65551 | NNNNNNNNNN | TAGTACTTAA | ACTGTAAAAA | GGAAATCATA | TTTTATAAGT |
| 65601 | GATGATGTAA | CATTGTTATA | ATTTGGCAAC | ACTTGTGTGT | TCTATTTTTA |
| 65651 | CAGATGATGC | ACTATTGTGC | TGAGAAAAGA | TTAATGTAAT | TTCCCTTTGG |
| 65701 | TCTTATGGAA | AGAAGCTCTG | GTTATATGGA | TAGGAGAGAA | CACCTGTTGT |
| 65751 | CTAGATACCC | CTAGTTCATA | GATTCATTTG | TTGTTTTTTG | TTGATTTTCT |
| 65801 | GATCCACTCA | TCTTCTGAT | TATTCCTTGG | TAGGTGTTCT | GGTACACAGT |
| 65851 | TGTAATTTTG | ACCATGGACT | TTGTGGATGG | ATCAGGGAGA | AAGACAATGA |
| 65901 | CTTGCACTGG | GAACCAATCA | GGGACCCAGC | AGGTAAAACC | ATTTCAATTA |
| 65951 | ACTTTTTCTG | GTACACATTT | CAATGTGATA | CTATCTGAAG | ACTCCACTGC |

FIGURE 3V

66001 TGCTAATCAA GTCTACTGTA ACACATCTCT GTGTTTACTT GATAAAGATG
66051 GTCAAAAAGG ACCATGCCTA GCATTTATAT TTCATTTCTT TGAAAGAGTT
66101 AAATTTGTGA GCTATCAACT AGCACTATAT TATGGGAGAC AAGTAGTTAA
66151 TTAAAAGGTC AACTAATTTT TCTCAAAAAG CTTGATAAAT AAAATATTTA
66201 GATTAATTCT TGCAAATAAT TCTTAAGTTA TTTGTAGCAC CCATTCCCAG
66251 GAATAAAAGA AGTAATATAG TGGTATATCA GCCATGGTAA TAGGCATTCC
66301 CCAGTTGNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNAGAT
66351 TTGAGTGGTA GTAGAAGAAC ACAGTGGAAG TTTCTATAAA GTTGTGATTA
66401 TCCTCTGAAA AAATTTCACT CTAACCTGGG TGCTGAAGGA AATGGTTACG
66451 AATCACCACC ACTCATTTTA TAACAAAAGG TAGATTTAGA CATAGTTCCT
66501 TGGAACTCAA AGGAGTTTTA AATTGCACAA TTCCTATTAA TCTAAAGTTT
66551 TAGTACTCAT AAGACATCTT TTCTCTAGA TTCAAGATAG CTTTTCTCCC
66601 CTCAGTAGTT AAATCNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66901 NNNNNNNNNN NNNNNNGAAG TTAAATCTTT TATGGATAAG GCCTACTTAA
66951 TGGAAGGGTG AGGAGTTTGG TTTCATTTTT TTTCTTTTTT ATACTCCTGA
67001 TTCTAGATTT TAAGAAGATA CAGAATGTGA AGGAAAAGCT TGTATTTCAT
67051 TTTATTTTAA AAGCCTAGTT CTTGAATACC ATACTGGATA AGAGACCAA
67101 ACACAGAAAA GTTAACCTGA GAGAGCTCAT GATATCTTAA ACCAGCTCNN
67151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
67201 NNNNNTCATG CATTCTAGC AGGCTCCCTG AAGATCTCTA TGGTGGTCCT
67251 CAACGCACAT TTTGAGTAAC AAGGTATTAA AAAATTAGTN NNNNNNNNNN
67301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
67351 NNNNNNNNNN NNNNATTTAC TTAATAATTA AAATGTGAAT TAAATTCCTT
67401 AAAATAGGAA TATCAATAA ATTCCAAATT GGTTCATACA TTTCTTTTCA
67451 TTGTAAAAGA CAATATCATT TAGTGCCTCA CACACAGTAC TCTCAGTCAT
67501 TCTTGCAATTA CTGTCTTACT CATTAAAGT TAACTCTGAG CAGGACAGCA
67551 TTATCATTAA TGTCACAGAT TCTGTGAATA AGAGAGAAAC CAACATATCT
67601 ACNNNNNNNN NNNNNNNNNN NNNNNTACAA TATAAAATAC CTATGACAGT
67651 CCTAGCTTAT TTAAATCTC ATCAATTTTA TTTAATTGTA GCAATTAATA
67701 TAAATTTATT TCATTGCTAC AGAGTTGTAG CAATTTACAC TCCTACCAAC
67751 AGTATTTGAA AACCATCCCC AACAGATATT ATAAAAGTAT TTTTAAAT
67801 CTCTACCAAT CTTATAGGTG AAAAGGAATC TCTTTTAATT TTAATTTGCA
67851 TTTCTTCTAT TGTGAAAAGT TGATCATGTT TTATCATTCA CTCGAGGAGC
67901 AAANNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
67951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
68001 NNNNCTCTGT ACTACTCATG CAGGCTCCTA GACAGTCCAA TACTCAGAAC
68051 ATCTGCAGCA TTATTGTGTT TTTGACTGTA CCACAGGCTG GCCTCAGTAG
68101 GTGTGCATGT GGTAGTGTGC ATGATCTGAT TAGCAGATCA CAGACCTGTA
68151 GATTAGACTG ACTGACCACT GACCTGGAGT AGTTTGCTAC AGAAGGAGGG
68201 ATCTTTGGCA GTCATGTATG ACAGAATGGT GAATAATATT TATGGTCATA
68251 ATGAATGATT CCTTCAAGTA AGTTACACTG GGGAGAGTAA TTCATGCTAA
68301 ACTCAGGAGA ATTCACTGCA CTCTATGTA AACAATCTAT GTAAAATACT
68351 TGAGTTTTAG AATTTAAATA CTGTATTTTA GAACCAATTT AGTCCTTTTC
68401 AACATTTTTT TATTATTCAT TATATTTGTA AATGTTTAAA TTTGTTGTGA
68451 CCCAAATCTC CTTTCTCTC TGATGGAAAA ATGAGAAGAG GGTAGAAAT
68501 TAGAATGGTA GACAAAGAAA TTTGAAATCG TTATTGGACC GTGTTAGTAC
68551 AGAGCAAATT GAAAGGAAGA GATGCCTGCT TGAGTGATTG GATAACTCGT
68601 GATGATTGGC TACGGTCCCT GGAGGATGCT TTTCTAGAAT TGGTCCAAGT
68651 TTTTACCAAG AACTTAGATA AATAGGACAT TGAACCTAAT ATTGAAGATA
68701 AACTGACCAA ACTACGGTTG ACTCAAATCT GGAGAGGAAA GTGACTATGT
68751 GTGCTGATAC AACCAGCATT CGAAAAATTA GATTAATGGG TCACAATTAA
68801 GAAGGTGAAA TGTAATGGGG TTGACTCGAA GTTATGGCTC ATAGTCGAGA
68851 AACCAGTCAC ACATCTGAGT GAGGGGGAAG ACCTGACATA TACTAAGGG
68901 TAGAAGACTT TAGAGTTTTA GTGTGTAATG AAATCAATCT AAGCGAACCT
68951 TGTGATGTTG CTGCCAGAAA TATGGACCAT CTCTATAGAA ATATGCTATG

FIGURE 3W

| | | | | | |
|-------|-------------|-------------|------------|-------------|-------------|
| 69001 | TTTCAAGAAT | AAAAATCCAC | TTTGGTTAAG | TGGTGAATCC | ACTTTTGAGA |
| 69051 | ATTCTTTTTT | TCTAATATGA | GAAGAATAGA | GTTGGTCACT | TAGAAGAGGA |
| 69101 | TGACCTGCAT | GGAAAGGTAT | CTCCAAGCTT | GTGAAAAACA | GGCAAAGGAA |
| 69151 | CCATAACGTT | TAGCCTGGAG | AAGAAATGAT | GCATCAGGGT | TCCTTTTCAA |
| 69201 | ATAATTGACT | AGGGTTAGAG | TTGTTCTCTG | TGACCGAAAG | GGGATGAGAA |
| 69251 | TAGGATTGAT | GGGTAAAAAA | TACAGGGAGA | TTCATTTTGG | CTCACTGTCA |
| 69301 | TCTAAAAAGA | GAACAGGCTG | ACATGGAAAG | AGTGAGTTAT | CTTGGTGGAA |
| 69351 | GCATTTGACT | GGGTGACATA | TGGCAAGATA | CTGTTAAAAG | GTTTTGAGCA |
| 69401 | TCAGATGAAA | GTTGGATTAG | AAAACCTCTA | AACCAGCTCT | GAGAGCCTGA |
| 69451 | TTCAGAATGG | AATGTGGAAA | TAGTCCCAAG | AAATTCAGTG | AGCAGGATTC |
| 69501 | TAGCTTGGA | AGAAGGATGG | GAAGACCAAA | AGATGGCAGA | GAGTTGGAAG |
| 69551 | GCAGGATATG | TGAGGATGGA | ATTGCAGAAT | ATATTTTTAG | AATATGTTGA |
| 69601 | GAGAATTATA | ACCCAGCACT | GATGGGAAAA | TACTAATAAA | GCTTACCTTT |
| 69651 | TCATGGT TTT | TTATTCTAGG | AAATTCAAGT | GGATGGANNN | NNNNNNNNNN |
| 69701 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNTC |
| 69751 | ACATGTGAAG | GCCAGAGAAG | AATATAAAAA | ATCNNNNNNNN | NNNNNNNNNN |
| 69801 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 69851 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 69901 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 69951 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 70001 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 70051 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | AATAATTTCT | GTGTATTATT |
| 70101 | CTAATTACCA | AGCTAATTAT | TTACCTATTT | CTATCTAGAT | TATACCACTT |
| 70151 | ATAAAAGAAT | ATGTGTTTGG | ATATCCAATT | ATGGATGATT | TTAATGATGT |
| 70201 | GCAGTTCTAA | TCCAAGATCC | AGATTTGTAA | ATATTCTTAT | ATGCTAATAT |
| 70251 | TTCTAAGAAA | ACTCTCAAAA | CTCAAAACCT | TAAGAAATAG | CTGCAAAATA |
| 70301 | AGTGCATTTG | CTAGCTGCCT | TTCATGGTGC | TATTAGGTTT | TATCACATTC |
| 70351 | AGTCACAGTG | AGATGTGAAA | TTAACAATGC | CTCTAAGAAA | TGGAAAATTG |
| 70401 | TGCTGCTGAC | AGACATGGTG | ATGCGCTGCA | GGTATGTAGC | TGTGGCCCTG |
| 70451 | GGGACACACC | CTTCTCCTGT | GTGTGTCAAA | GCTAGAACTA | AGGCCCTTTT |
| 70501 | CCTGGAACCT | TGATCTCTGG | AGAACAAGGA | TTAAGGGCAA | CTGACTCAGC |
| 70551 | ACTCTGTCTG | CCACTGATAG | CTACAAGGAA | TTTGCTTACA | TATGTATCAA |
| 70601 | GTATGCCAGC | ATGTACATAC | AAGTATGTCT | ATTTCAAGCAT | TAAAACATTA |
| 70651 | TTTTAAAATA | CACATAAATG | TAAAATAAAT | CAACAATATA | AGTGAAAGCT |
| 70701 | TTCTACCACC | CTGACTCCC | CATTCAACAC | CCAGCAAGAT | GCAGCCAGTT |
| 70751 | TCCTGATTAT | GTTCTCAGAG | AGATATTTTT | GGATACAATG | GCATTCTCTT |
| 70801 | CTGTGCCTTG | CCTTTGTAC | TTATCGATAC | TACCTTAAGA | GTTAACAGGN |
| 70851 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 70901 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNTGGGG | ATAATAAGTA |
| 70951 | CTATCCACTA | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71001 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNGC | TGAAGCATGT | CCCCTGTCAG |
| 71051 | TACATACAGA | ACACNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71101 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71151 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 71451 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNTCAT |
| 71501 | TTTTTCTCTG | ATGTTCTTTA | TTTTTCTTTA | GTATTTTAT | TGATTTATGG |
| 71551 | AAGTTTCTA | TACAAAAAGA | AAAATTATCC | TTTTTTAATT | TGTTGCAAAT |
| 71601 | ATTTAAAAAC | ATTTTGT TTT | GTCTATTTTA | TGGTATTTTT | ATCTTAACAT |
| 71651 | GCAAATTTGG | AACTTAAAT | GTCTTCAGT | AAGAAATTGG | CTAAATGGTT |
| 71701 | ATTGGCATAC | TTGCAGCAAT | GCAAAGTGTA | TTTGCTACCA | TAGAAAAATG |
| 71751 | GTTATACAAG | AAGCAACTTA | CAACACTGCA | AATATAGTAT | TCCTTTTTTTT |
| 71801 | AAAAAAAGTT | CATATTCATG | TGCATTTATG | TGAATGTATT | ACATAAGAAA |
| 71851 | AGTTCAGAAG | TCAAACCTT | CACAAATACC | TCTAGATTTT | GTGAATTGAA |
| 71901 | AATTTAAAGG | CACCAACCAC | TACAGAATTA | AAAATAAATA | AACTAACATT |
| 71951 | TTATAACCCA | ACTTCTGGT | CAAGTTTTCG | AGCATCTATT | ATTTACCCTC |

FIGURE 3X

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

| | | | | | |
|-------|------------|------------|------------|------------|------------|
| 72001 | CCTCCCCAAA | ATTCTTGAGA | TATTGTGATA | CTGTAAACTG | AATGAATGTT |
| 72051 | TAATATAATT | GTACATTTAT | CACCAGAATT | ATGAAATCAG | AAGTGAAAAT |
| 72101 | TTTAATAGGT | AACCTGCTTG | ACACTGAATC | AATTTTTCCT | TTAGTTAATG |
| 72151 | TTCCAACCTG | TATCATTTCC | TTTATTACAA | AGGGAACGAA | TTCTTCATGG |
| 72201 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 72251 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 72301 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 72351 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN |
| 72401 | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNNNN | NNNNNNNTAC |
| 72451 | CTGTGATGAT | TCAAAGCTG | ACTTATAGTG | AGACATGTAA | ATATTTTGGC |
| 72501 | CCAAAATATT | GTTTGGATTA | GTTTGTGGAG | GAAGGGGACA | AGGGAAGTGG |
| 72551 | GAGAAAGATT | CTTAAGACAT | TATGTATTTT | TAGTGATTCA | AGTACGTGAA |
| 72601 | TCAAGGATCT | ATTTCCCTGC | AGTTTTTTGC | TGTATCCTTG | GAACATCATT |
| 72651 | CGTTGATTCA | AACATCCTTA | AACACCTGCT | GCCTGCTAGC | AGGGTAACAC |
| 72701 | AGAGGTTTGG | GGGAACTTT | GGTTCAGACT | TTTAAGTTGT | TCACAACCTG |
| 72751 | CTGGGGAGCC | CTTAGCACAG | AGGCTGGCAC | ACAGTAAGTG | CACAGAAATC |
| 72801 | ACCTGTGACC | TATAGAGCTC | CGACGGGACT | TGTTAGAGAA | GACACTGGAA |
| 72851 | TAAAAAGTTA | ATACCAAATT | TTATCATGTC | CTGTATAAAT | TAAGCATAAG |
| 72901 | AAAGAGATGA | TAGTTTATAT | CTGTTAACTG | AAAACATAAT | GGGATTACTA |
| 72951 | CCAAAATAGA | TACACATGTT | CAACATTCCA | ACCATTTTTC | TATGGAAAAA |
| 73001 | CCAGTGGTTG | AACATACAGT | GTGTTTTTCT | TTGACTGAGA | TAAATATGCC |
| 73051 | TGAGGGAATG | TCATTTGTAC | ATGGATCAAG | TTTCAAAATA | CTACTTGCAA |
| 73101 | CTTCTCTCT | CTTGAGAAGC | AGCACCCCCC | ACTTCACTGC | TGCTTAACCA |
| 73151 | TTTTTTTTCT | AATGCAGTTT | CCTTAACAAC | AGAGGGAAGA | AAACTCTGTA |
| 73201 | GCCCTAGGGA | TCTGAGAACA | TCATCAAAGT | GGTCTATCTA | AAGTGAAAGT |
| 73251 | TTTTATTAAT | ATTTGTGTTT | ATGTATATAT | AATAGGACAT | ATTTATAGAA |
| 73301 | TAGATATGAT | ATAGATTTGT | TTTACTATAA | AAATGTTTTA | AATGGCTTAC |
| 73351 | CTCCAAATAA | AATTGAAGCT | CCAGGAAGAC | GGGGAGGGTT | TCCTGAATAC |
| 73401 | CTCTGTATCT | ACCTGGTGTT | AGATTTACAC | CAGTGTGAGA | AGCTCTGCCA |
| 73451 | ACTCTGATTT | TTTTTTTAAA | GAACCTTTAC | GTTTAGAGAT | TTTTTTTTTT |
| 73501 | TAAGTTGGGG | GAGGAGAGAA | TAAGCAGATA | TAGGGCCCTT | TAAATCACCA |
| 73551 | TCTGCTTTAC | TGTTTCAGAG | TTAATTATCC | ACCAGAAAAT | TCTCCCTAGA |
| 73601 | AATTTAGGGT | CAGTTTTACC | AAACCACCAA | GCCAGATCGC | CAGGGGTGTT |
| 73651 | GGACAACGTT | ATCATAAACT | GGTGAGGGTC | ACAAGGTGGT | GAATGTGTAT |
| 73701 | ATGTTGTCTG | ATGCTATAT | GTCTGATCAT | ATTCCATTGG | AAATATTCAT |
| 73751 | AGTGTTTGGT | TACTAATTAT | TGTGTGTCTC | TGGCATTACT | TGTACTTTCC |
| 73801 | ATGAGCAAAG | TAAGTGAATG | TTTTCTGTAA | ACACATACCA | TAGAACTTAA |
| 73851 | GACAGGCATA | CAAATAAGTC | ATATTTTTCT | TCTTCAAAA | TCTGGGAATT |
| 73901 | CTTTACAAAG | AAAAGAAAAA | AAAAAAAAAA | CTAAACTCC | TGTCCATGCT |
| 73951 | GCTTGTTCTA | GAAAGGGATA | TTGGCACATA | CACACAGCCC | TGCTTTTCCC |
| 74001 | ATTCCAGGTG | GACAATATCT | GACAGTGTCT | GCAGCCAAAG | CCCCAGGGGG |
| 74051 | AAAAGCTGCA | CGCTTGGTGC | TACCTCTCGG | CCGCCTCATG | CATTACGGGG |
| 74101 | ACCTGTGCCT | GTCATTCAGG | CACAAGGTGA | CGGGGCTGCA | CTCTGGCACA |
| 74151 | CTCCAGGTGT | TTGTGAGAAA | ACACGGTGCC | CACGGAGCAG | CCCTGTGGGG |
| 74201 | AAGAAATGGT | GGCCATGGCT | GGAGGCAAAC | ACAGATCACC | TTGCGAGGGG |
| 74251 | CTGACATCAA | GAGCGTAAGT | AGATCCACAA | AGGAGGCAGG | ACCTGGGACG |
| 74301 | TTTTCTTTTC | ATAGGAGAAC | TCTGGGATCT | GAATTTGAAG | AAGCCTTGCT |
| 74351 | GTGTGAATTC | AGGCTCAGAT | AAAGGTTTGG | GTTTTTTTCT | GGGCCTGATG |
| 74401 | ACTCCATTCA | GTGTCTCTTT | CTGAAACCAC | TTTCTGCTTC | TTCATCCTTC |
| 74451 | TTCAATTGCA | CCTGTGATCA | CCACCTCTCC | CTAACCTTGG | TAAAGATCCA |
| 74501 | TTTTTTTTTT | ATGGAAAAAA | AAATCTAGGA | ATAGAAAAGT | TTTTTTGGTT |
| 74551 | TTTTTTAAAG | TTGATCAACA | GAAACTTCCA | AGGGTGGGGT | GGTGAGCTTA |
| 74601 | ATTGTTCTCT | CTTTCCCTTT | CCTTTTAATT | CCTTAGATTT | CAAACCTTTA |
| 74651 | CTACCATAAA | TTACCCAGGA | GATGGTTCTT | TTTTCTTTTT | TGAAATGTTT |
| 74701 | CTTTCAAGTT | CAAATTTTTT | TCTTTGTTTC | AGAGATCACA | GTTAAGCAGC |
| 74751 | GTAGGGTGGG | AACTCAGAAC | TACAATTGGA | AAGCACTATT | CTATTCAGGA |
| 74801 | CAGGATGTGG | GAAGTGCTTT | GATGGAATAT | ATAAATTTGA | TCTTAAGTAA |
| 74851 | TCAAGACATA | GGTAGCAAAA | AAATGGTCTA | GGGCATCCCA | ACAAAATTAT |
| 74901 | GAATGCATTT | GGGAAAATGT | CATCATTTTG | CTGGTGCTTA | TGAAACTTGC |
| 74951 | TGTTGGGAGA | TACCTATGCC | TGTATGTAGG | TATAGTTTAT | AAGAAACCAT |

FIGURE 3Y

75001 TTTAGAAAAGA AAAGATAGAT TCCATTAATA CAGTTTTTAT CAGAGTACTT
75051 GGATTTTGTT TAATTCTTAC ATATTTTTTC TTAAAACCTT TCTCAGTATT
75101 TTTATTGTTT AGAGAAATAA AACAAGATAA TCATTTTAAA TCATAGCACT
75151 TACGTTTTCT CTTGTTTTAT AAGGAGCAAG GATGCTCTAT AGAAAATATA
75201 ATGTAAGAAT AATAAAAGTT TTTGGTTTTT ACATAGGTAA AGCAACAGTG
75251 TGATTGGATT ATGGTGTGTTG ATTCTATTCC ATTTTCAGCA AGAAAGCGAT
75301 GTTAACCACA AAGGAACTAA GAAACATTTA AGATAGGCTG TGTGATTATG
75351 ATCTTTTCAG TCTTTGGCTC CTAATATCTG TTCCTTTATA TTCTATCACA
75401 CTCTTCTAAC TTTGGTAATC CTTGACAAAA GTGTGCACTT TATAAACAAAT
75451 CCTAAATCGA ATTGGTCTAT AGCTTAGAAT GGCTTTTTAA AGAATAATTG
75501 ATTCTGAGTA ATGTGGTCTG ATGAACAGTT TGATGACTTC AGTTTCTACT
75551 GAAAGAGAAG CTTCACTGAT ACTGGCAACT ATATCTGTT TTTTTCCTCC
75601 TGCAAAATAA GTTTAAATTT GGTTTGGGGG AAGGTTTGCC TTTATTTTTG
75651 CTTAATAAGG AGGCATTAGA AAGGGGCAGA GGAGGCTTGA CTGGTGTGTG
75701 CTCTTCTCC CTAGGTCGTC TTCAAAGGTG AAAAAAGGCG TGGTCACACT
75751 GGGGAGATTG GATTAGATGA TGTGAGCTTG AAAAAAGGCC ACTGCTCTGA
75801 AGAACGCTAA CAACTCCAGA ACTAACAATG AACTCCTATG TTGCTCTATC
75851 CTCTTTTTCC AATTCTCATC TTCTCTCCTC TTCTCCCTTT TATCAGGCCT
75901 AGGAGAAGAG TGGGTCAGTG GGTGAGAAGG AAGTCTATTT GGTGACCCAG
75951 GTTTTTCTGG CCTGCTTTTG TGCAATCCCA ATGAACAGTG ATACCCTCCT
76001 TGAAATACAG GGGCATCGCA GACACATCAA AGCCATCTGT GGGTGTGCGC
76051 TTCCATCCTG TGTCTCTTTC AGGAAGGCAT TCAGCATGCG TGAGCCATAC
76101 CATCCTCCAT CCTGATTACA AGGTGCTCCT TGTAGCAAAT TATGAGAGTG
76151 AGTTACGGGA GCAGTTTTTA AAAGAAATCT TTGCAGATGG CTATGATGTT
76201 ATGTGTTCCG TGTTGTACCA TGAGTAGTAT TGACTTCCCT TGAGATATGA
76251 TGTACAATGT GCTTGTGAAA TTGACTTACC CTCTTCACTT AAGTTAGTTC
76301 TGGCCTGACC TGAAGTCTGA CTTTTACTGC CATTCACCTT ATAAAATAAG
76351 GGTGTGTAAC ATATCAAGAT ACATTTATTT TTATCTGTTT TTTTTTCTC
76401 GTTAAAGACA ATTATGTAGA GTGGGCACGT AATCCCTCCT TAGTAGTATT
76451 GTGTTTTGTG TAAATGTGCT ATTGATATTA AGTATTTACA TGTTCCAAAT
76501 ATTTACAGAC TCTAGTTGCA AGGTAAAGGG CAGCTTGTTG TCTCAAAAAA
76551 ATACATGGTG AAATGTCATC CAGTTCCATG ACCTTATATT GGCAGCAGTA
76601 GGAAATTGGC AGAAGTGTG GGTGTGTTG ACGGAGTGAT GAATTTTTTT
76651 TTAATGGCCT TGAGTTTGAT CTCTGCAAAG GATAGGAAAC CTTTAGGAAG
76701 ACAAGAAACT GCAGTTAATT TAGAACTGTC ACTGTTTCAA GTTACACTTT
76751 AAAACACAG CTTTTACCAT CATAACATGG CTCTGGTAAT ATGTAGGAAG
76801 CTTTATAAAA GTTTTGGTTG ATTCAGAAAA AGGATCCTGT TGCAGAGTGA
76851 GAGGAAGCAT AGGGGGAAC TCCATTGGAA CAGATTTTCA CACAACGTTT
76901 TAAATTGATA TAAGTTTAGG CAGTTGTAGT TCATAACTTA TGTTGCTCAT
76951 GTTGTGCTGT GTCAGGATGG GATAGGAAGC AAGTCCCATG CTTAGAGGCA
77001 TGGGATGTGT TGGAAACGGA TTTACACACA CTGGAGGAGC AGGGCAAGTT
77051 GGAATTCTAA GATCCATGAA CCCCCAATG TATTTCTCCT CTGCATATTT
77101 TACCAATATA TTAaaaaaca ATGTAACCTT TAAAAGGCAT CATTCCTGAG
77151 GTTTGTCTTA ATTTCTGATT AAGTAATCAG AATATTTTCT GCTGTTTTTG
77201 CCAGGAATCA CAAAGATGAT TAAAGGGTTG GAAAAAAGA TCTATGATGG
77251 AAAATTAAAG GAACTGGGAT TATTGAGCCT GGAGAAGAGA AGACTGAGGG
77301 GCAAACCATT GATGGTTTTT AAGTATATGA AGGGTTGGCA CAGAGAGGGT
77351 GGCGACCAGC TGTCTCCAT ATGCACTAAG AATAGAACAA GAGGAACTG
77401 GCTTAGACTA GAGTATAAGG GAGCATTCTT TGGCAGGGGC CATTGTTAGA
77451 ATACTTCATA AAAAAAGAAG TGTGAAATC TCAGTATCTC TCTCTCTTTC
77501 TAAAAAATTA GATAAAATTT TGTCTATTTA AGATGGTTAA AGATGTTCTT
77551 ACCCAAGGAA AAGTAACAAA TTATAGAATT TCCCAAAAGA TGTTTTGATC
77601 CTAAGTAGTAT TATGCAGTGA AAATCTTTAG AACTAAATAA TTTGGACAAG
77651 GCTTAATTTA GGCATTTCCC TCTTGACCTC CTAATGGAGA GGGATTGAAA
77701 GGGGAAGAGC CCACCAATG CTGAGCTCAC TGAAATATCT CTCCTTATG
77751 GCAATCCTAG CAGTATTTAA GAAAAAAGGA AACTATTTAT TCCAAATGAG
77801 AGTATGATGG ACAGATATTT TAGTATCTCA GTAATGTCCT AGTGTGGCGG
77851 TGGTTTTCAA TGTTCTTCA TGGTAAAGGT ATAAGCCTTT CATTTGTTCA
77901 ATGGATGATG TTTCAGATTT TTTTTTTTTT AAGAGATCCT TCAAGGAACA
77951 CAGTTCAGAG AGATTTTCAT CGGGTGCATT CTCTCTGCTT CGTGTGTGAC

FIGURE 3Z

```

78001 AAGTTATCTT GGCTGCTGAG AAAGAGTGCC CTGCCCCACA CCGGCAGACC
78051 TTTCC'ITCAC CTCATCAGTA TGATTCAGTT TCTCTTATCA ATTGGACTCT
78101 CCCAGGTTCC ACAGAACAGT AATATTTTTT GAACAATAGG TACAATAGAA
78151 GGTCTTCTGT CATTTAACCT GGTAAAGGCA GGGCTGGAGG GGGAAAATAA
78201 ATCATTAAGC CTTTGAGTAA CGGCAGAATA TATGGCTGTA GATCCATTTT
78251 TAATGGTTCA TTTCC'ITAT GGTCATATAA CTGCACAGCT GAAGATGAAA
78301 GGGGAAAATA AATGAAAATT TTA'CTTTTCG ATGCCAATGA TACATTGCAC
78351 TAACTGATG GAAGAAGTTA TCCAAAGTAC TGTATAACAT CTTGTTTATT
78401 ATTTAATGTT TTCTAAAATA AAAAATGTTA GTGGTTTTCC AAATGGCCTA
78451 ATAAAAACAA TTATTTGTAA ATAAAAACAC TGTAGTAAT ACCAGTTGTC
78501 TATTCTTGTT TTTTGAGTTT TGTTTTTTTT TGA'CTTGGA AAAAGCATTG
78551 AGGTAGTTAA ATGATGTTTC ACAAAGTCA TAGTAGAATC CCTTTTACTG
78601 TTTGGATGGT GGAACAAAG ATGTTGCCTG CAGTATTATA CTTTCTAGGT
78651 TATAAACAT GAGACACTTT ATTTTTTTTA TCAGCATGAA CAGGGAAAGA
78701 GATCAGAAGA TCACTATAAC CCATGCCATG CCTTAGTAAA TTGCTTTAGT
78751 TATGTTTAT TATCATTTCA TTGTAAACAT TTGCT (SEQ ID NO:3)

```

FEATURES:

Genewise results:

```

Start: 3000
Exon: 3000-3070
Exon: 5248-5348
Exon: 19533-19583
Exon: 34426-34518
Exon: 39701-39790
Exon: 44102-44221
Exon: 45394-45513
Exon: 47168-47302
Exon: 47607-47729
Exon: 49126-49127
Exon: 49132-49522
Exon: 65277-65363
Exon: 65834-65932
Exon: 74008-74264
Exon: 75715-75807
Stop: 75808

```

CHROMOSOME MAP POSITION:

chromosome 4

ALLELIC VARIANTS (SNPs):

| DNA | | | | Protein | | |
|----------|-------|-------|--------------|----------|-------|-------|
| Position | Major | Minor | Domain | Position | Major | Minor |
| 3365 | C | T | Intron | | | |
| 3411 | G | A | Intron | | | |
| 5243 | G | A | Intron | | | |
| 6532 | T | G | Intron | | | |
| 9208 | - | G | Intron | | | |
| 15833 | T | C | Intron | | | |
| 27795 | G | T | Intron | | | |
| 31341 | A | G | Intron | | | |
| 33370 | G | A | Intron | | | |
| 37788 | G | A | Intron | | | |
| 41465 | C | T | Intron | | | |
| 41466 | A | G | Intron | | | |
| 41653 | T | G | Intron | | | |
| 47666 | G | A | Exon, coding | 281 | V | I |
| 52613 | G | A | Intron | | | |

FIGURE 3AA

| | | | |
|-------|---|---|--------|
| 52645 | G | A | Intron |
| 59197 | - | A | Intron |
| 63508 | T | C | Intron |
| 75153 | C | G | Intron |
| 76017 | C | G | Intron |
| 76033 | C | T | Intron |
| 77194 | A | G | Intron |
| 77508 | T | A | Intron |
| 77557 | G | A | Intron |

Context:

DNA

Position

3365

CGGGAGGTGAGCTGGGCCCCGGGGCGCCCTCTCCTCCTTCCCGCGCTAATTTCACTCA
CTGTCTTGGGTCACTTTTCCCGCGGGTTTCGTGGTCAGAGAGGCGTCTCCTCCATCCA
GAAGTTGGGCCACCGCACAGCGTGGCGCGAGGAGAGCGGTCCAGCGGCTCCGAGTGCCCCG
CCCGAGGCGGAGAGGGCGCGCCCTTGCGAGTCTGGGACCCCATCCGCGCCCCCGAGGG
CGACTCGCCCCGGCTCGGGAATTAGGACTGAGGGAGAGAGCCGCTGGAGCCTGGGATCT
[C, T]

GGCTCTGAGGGCGCGGTTTAGCCACCTACGCCGAGGTGACGCGGAAACATCCCTTACCC
GGGAACTCCCGCGCCTGAAGTACGAGGCTCTTCACTGGGGAAGCTTCCAGGCCCGGG
GGGAGGCCCGGGCTCTGCTCAGGGCTCTCGGGGCGCTCACACAGAGAGTGGGTGCGAGT
CAGCGACTGGGCTACGGGGGAGATTTGTGGGCCTCTCCATTTGGTTTTCTTGAGGGAAGG
AGACTCAAAATGAGGACCGGAGGGTGGGCGCTCCGTGAATGTGAGCATGAGTGTGTGGAT

3411

TAATTTCACTCACTGTCTTGGGTCACTTTTCCCGCGGGGTTTCGTGGTCAGAGAGGC
GTCTCCTCCATCCAGAAGTTGGGCCACCGCACAGCGTGGCGCGAGGAGAGCGGTCCAGCG
GCTCCGAGTGCCCGCCCGAGGCGGAGAGGGCGCGCCCTTGCGAGTCTGGGACCCCATCCG
CGGCCCCCGAGGGGCACTCGCCCCGGCTCGGGAATTAGGACTGAGGGAGAGGAGCCGCT
GGAGCCTGGGATCTCGGCTCTGAGGGCGCGGTTTAGCCACCTACGCCGAGGTGACGCGCG
[G, A]

AACATCCCTTACCCGGGAACTCCCGCGCCTGAAGTACGAGGCTCTTCACTGGGGAAGCT
TCCAGGCCCGGGGGGAGGCCCGGGCTCTGCTCAGGGCTCTCGGGGCGCTCACACAGA
GAGTGGGTGCGAGTCAGCGACTGGGCTACGGGGGAGATTTGTGGGCCTCTCCATTTGGTT
TTCTTGAGGGAAGGAGACTCAAAATGAGGACCGGAGGGTGGGCGCTCCGTGAATGTGAGC
ATGAGTGTGTGGATGTGTGTGTGAGAGCGCACACACTGCGCCGCTCCTCAGACTCGGGCG

5243

GGCAGCCTTGAGAATTGTTGGTGTTAGCTGGTGCTATTTCTACTTAAGAAAGCATCTTTC
CCTCCCGTTTTTTCTCCTACCCCTTGTGTGTTGGGGGATGGGGTATTTAACTGTCTGAAAT
TTGACATTAGATCATAGCAGAACTATTTCTGTGAAGGTGTTTTAACCTTAAATTTCTGCA
GGCATAAAAGAGTTTGTAGAAAATTGTTGGGTGTGGTTGACATTTTTTCTGAAGTAATA
TACAGAGGAAATTACCTTTCTCTCAAAAGTGTCTTATTTATTTTGAATCTTTTTTTTT
[G, A]

GTAGGTGGCCAGGCAAATAGTGTATCGATTGGCCTATGTCGTTATGGTGGGAGGATTG
ACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTGAGCGTGAGTATCAAGCCT
GGGGACTTCAGTTCCTTGGGAGGTGTGGCTTTCCACCTTGTTTATGGCTTACCCCCACAT
ATCAGAGGGTTCACTAGCAAGGCTTGGCCTTGCAGGTCTGACTTGGGGATTTTCAG
GTACAGTCCAGACTCCTTATCTGCTTCTTTTACAGCTTTAGCCACCTGTATTACGGCCCA

6532

GGGCGGGGGTAGGGGGCACTTCTGGTAAGATGATGGGAACTAAGTTGGGTCTACATTGG
GATATATATTTTATTGCTAATGAGGAGGAGGCTTAGAGGAAGAGAGAAGGGCAGTTACGA
AGGCTAGAGCTGGCAATGGAGAAGCCTGCCTTAGAGATGGGTTGCTAGTGTGAGGAGTCA
GGCAAATTTAAGTTTCAAGAAAGTTAGGAGTTCCCTCTGCTATTTTAAATTTTGGAGGATGC
TTGCAATGTCTTCCTTAATTTTGTGAAAGAGGGACAGTGACAGTCACAGATTGACTCTAA
[T, G]

TGCACATAAAGACCACAATCTCTGGTTGGGAATAGAAAGGTAAGGAAATGAATGTTTGC
CTACCTGGTATGGAATTTGAGAACCAACAGATTCTAATAACCAAAATGTGAAGAAAGGAC
CCTTCTGTTGGCCCAACACACCTACACATAACCCCTCCTGAGTGAAAAATGAGTAGTTCTA
TACCTGCAGTCTCCAGCTGTGCAAATACTTCTGATACTACAGAAGACTAAATTCACCAG

FIGURE 3BB

GCACCATTCTTCTTTTTTGATCATCTTCCCTTAAACAATATTGAATAGACTAACCAGTG

9208 TCTGAGTTCCTTTGGCTGTCACTTGAGCAACTTGTGGTTTCTGAAGGAAATTATGTGAATTA
GGATGGTTTTGTATCATTTATCCTTAAGAACAGGGAAAATTGAGATGTTTTCTTATGTTT
CTGCTGGAGATTTTGGAAAGATGTGAAACCTACACCTACAGATTGACCTTGCTTAGTTAG
CTCTGAACCTCCTGCTGCCTCTTCCACGTAAAGTGAAAATTTGGATTCTTATCGGCTTC
AGATAAACTTACAGGTTAGTGAAACATAGGGACTGAGATATAGTAATTCATTCTGAAGCT
[-, G]
TTTTGGAGTGGTCAAATAATTTTAGTTGGATAATATATATTTACTGGCTAATGATTGTGGA
TATTGGAAGTGATGAAAAAATTATTGAATTATTTCTTCTGCATTTCAAATGAAAAGGCT
ATTAGTTTGAGCAGAGAATTTGATTTAGTAAACAAAATATTTAAATTTTCATGTTTCATT
TCTTCTCCTATCTGGGTTGAGATACTCAGTCTTATAAATGGAACATGATTATTTTTGTC
TCCCTAACTGGTTATTAACCTCCTGTCCATAATCACAAAATATATAGATTATATATTT

15833 CATACATTAGCTTTGGTAGATACTAGATGCCAGGGGAGTGCAAATTTGAGAATAAGGATT
TGGCATGGGTATTTGCATGTTGAGAGTCAAGATTTAACTAAATTTTAAAACCAAACACT
TCATTTGATCAATATCTTTTAACCATTCGTAGATTAATAATATAAATCTCCAGAGCT
GACATTACTTTGCTATAACATCATCAGATCACAAGATTAGGGTTGCCTTTTGTAGATGTT
ATTCATTACTCTGAAATGTAGGAAATATGTCATCACAGTTTAAATTTGTAGTAATATA
[T, C]
ACAAAAGGAAAAACAACTAGGAATTTTGGATTTATGCTTACTTTGCCAAAAACCATGTT
GATTTTCAAAAACCTTTAGCC

27795 ATGAAAACAGAAAGAGAAAGTTAAATTTCAAGTAACTGAAAACAACAATAGTGCTCAG
AGGATTTTTTTTAAAGTAAAAGGGATAGTGCTTAATTATGACAAATAAAAGCTAACTTG
AGATGCACATATACAGATGCAGCCACTTATTTTGGCGGGGACACTTCAGGAATTAATAAT
TTAAATAGCGAGAGATCAATAGATACTTGGTAAATGTGTCTGAATTGGATGTTCCCGAGA
CACAAAAATAAAATGAGTTATTGACAGCTCTTGGGAGACAACATTATAAAGACTAGACAC
[G, T]
TTATTTATTTTAACTCTATGTTCTAAATTACCATTGAGTAATTGACATTCGTATTTGACT
ATGGTTTGTGGTTAAGTTCTTAATTGCAATAATGTTAAATAAAATGTGAAGCCCAAAGCA
AACAACAACAAAAATTATAGCAATACTTCAACAGAGGTAATAATAATATGCTGCATCAAT
GGTTTCAGAAATCCAGCATCTACATAAAACAAGCAACAGGGTAATGAAATTATTTCTTTTC
AAATATTTCTGGCAGAGCTACTTTAGTTTTCTTAAGTTATAGATTGTGGTCTTAACTGCAA

31341 CCACTTTGCTTATATTCTGAAAGTCTTTTGTTCCTATTCCACTGCTCTTACCTCTAACA
CACTGCCTCTAACACACCAACCTGCAGTTGTAGTTATTACACACCCTCCTTGGTTCTTTTC
ATCTCTCTATCACAGCCCTTGTGTGGTTTAGCCAATATATTTTAGTTCCACAGCTAAAT
TTTCATACCCTCTATGACTCTTAATCCCCTGCCACACTTGCTTACTATAATACATTATA
TATATAACAAATGTTTGATACGTATTTATTGAATTCATTCCAGAACTAATGCCAGCAAG
[A, G]
TAACCTTTGTGCTATATAGGAGAATATCTTTTGTGCAACAGTTTCAAAGGGTTTTCTTT
TTCTAAGAAGAAAGAAATTGATTGTATCAACTTTATGAGTATCCTACCGCATTTAATAGC
CATTGGCTAATCTAAGGGTTCCTGGTTACTTCACTGAATAGCCTATCAGATGGAAGTGCA
AACAACAGTTTGTGTTTGAATAGGACTCCCTAAACATGGAAGAAACATTAAACAGTGTGG
CCTGTTGGAATGTGTGCATTGTGCTCAAGATTAGGGCACTCTGCTTGAGAACAAAT

33370 CCACTCACCTGATTGTTTCTCACTCTTCCCTCCACCATTACCCATCTCTTTAGCTTAATG
AGTCCCCTGTGATCTCCCACTAAACAGCGGCTTACTTGCTGTGAAATATTCTTCTCT
TGGGTAGTCTGCTCCCTTCTCTGTCTACTCATGCTTCAAGATTCAACATAAGCCTCCTCT
ATGAGGCTTTCTGCACGTATGTATATGGATTGCTTGTGTAATGATTTCTTCACAGATTT
CATATTGCTGATAAATAAATATTGTTTTGAATAAGAAACGTGGTTTTGTATTTTTATCTC
[G, A]
ATTGTAGACTCCTTGAGACCAGTACCATGCTATACAATTATTTTTCATCTATTATAGTGT
CTGGCATAGGGACATGCACATATTTGGTACAGAA

37788 GCCCTGTCAGGATCAGGGAAGCTGCCATGTGTGGCCTGGTTGAGTGTGAGAGCTGCCTAG
AGACTTCCATAAAAAGTTGTTAGAGAAAATTGTTGATAGTGCTACATAGCAGATTAACTT
AAACTGTTTTTCATTACAGGCCACACATTTAATAAAGTAGAAAATATGCTTCACAGATAA
GGGAAATCAACAGGCTCCTTTTTTCTGGAGGAGAGAAATGTCAAAAAGAAATTAAATTTG

FIGURE 3CC

AAATAACTTTACAGAACTGGAAATTAGCTTTTGATTAAAAGTAGCTTTTGGTATATGACA
[G, A]
GTATTCACTGAGAATTTTGTAGCGAGTTATATACTTTAAGAAATAACCCCCAGAACTTG
CATCATGGTGTAAACAGCTTGAATAAACAAGTGCTTAACCACTGCCTTTAGAGCTGCCTG
GGAAACAGCCAGAATACCAGGGCAAGCTGCATTTTGGAACTGGTTAATTTAGTAGCCTT
GCCACAGGCTTAGTGTGATCTGCTTTTGGTGGCTTGATCTTCCCCACTAAGTCATTTTCT
GGATTTGTTACACCTAGAACTGTTAGGAAATTACAGGCTTGGGCTGATCATTAACATACT

41465 TTTATTTCTAGGTTCTTTTATCAACTTTTAAGATTTATTTGAGACAGTATGATCAATGAC
TTCATTTTGTCTGCTTATTTATGTAGGAGTATTTACTATAATTTGGAAGTAATTTATTT
TTGAATTTATTTGCTTAATTGAATGATCTCCAATAGATTGTGATAATGAACACAGCATTTA
TAGAAAGCAGCACATATTAACCTTACTTAATATGGCACTAGGTCAATGAGAAAAGAAGGTA
ACATAATTGAAGACAAGAACTCTTAAGAAAACCTGAGGACAAAAAGGCTTCTCACCAGGA
[C, T]
ACCAGATGCATTTAATCTTTTGAAGCTCTGTACTTTAGGAAAAGTCTGATATTTGGCAAA
TTTTGATAAACATGGATGACTATGGAATCCTATTTTATAGTATCTGAAGTGGCTTTCATA
AGGGTCATTGTGAAGTTTTTAGGAGACACCTGCCTGTGGCAGATGGGACAATGATGGCAG
TCACTAGTGATATTAACACCAGTCAGCTGTCAGGGAATATCATCCAGACCATCAGCAGCT
GGTAGAGTACAGCTTTCTCAATTGCTTTCCATGTTTTGGATACTTATATGCCCGTTAATA

41466 TTTATTTCTAGGTTCTTTTATCAACTTTTAAGATTTATTTGAGACAGTATGATCAATGACT
TCATTTTGTCTGCTTATTTATGTAGGAGTATTTACTATAATTTGGAAGTAATTTATTTT
TGAATTTATTTGCTTAATTGAATGATCTCCAATAGATTGTGATAATGAACACAGCATTTAT
AGAAAGCAGCACATATTAACCTTACTTAATATGGCACTAGGTCAATGAGAAAAGAAGGTA
CATAATTGAAGACAAGAACTCTTAAGAAAACCTGAGGACAAAAAGGCTTCTCACCAGGAC
[A, G]
CCAGATGCATTTAATCTTTTGAAGCTCTGTACTTTAGGAAAAGTCTGATATTTGGCAAT
TTTGATAAACATGGATGACTATGGAATCCTATTTTATAGTATCTGAAGTGGCTTTCATA
GGGTCAATTGTGAAGTTTTTAGGAGACACCTGCCTGTGGCAGATGGGACAATGATGGCAGT
CACTAGTGATATTAACACCAGTCAGCTGTCAGGGAATATCATCCAGACCATCAGCAGCTG
GTAGAGTACAGCTTTCTCAATTGCTTTCCATGTTTTGGATACTTATATGCCCGTTAATA

41653 AGCACATATTAACCTTACTTAATATGGCACTAGGTCAATGAGAAAAGAAGGTAACATAATT
GAAGACAAGAACTCTTAAGAAAACCTGAGGACAAAAAGGCTTCTCACCAGGACACCAGAT
GCATTTAATCTTTTGAAGCTCTGTACTTTAGGAAAAGTCTGATATTTGGCAATTTTGAT
AAACATGGATGACTATGGAATCCTATTTTATAGTATCTGAAGTGGCTTTCATAAGGGTCA
TTGTGAAGTTTTTAGGAGACACCTGCCTGTGGCAGATGGGACAATGATGGCAGTCACTAG
[T, G]
GATATTAACACCAGTCAGCTGTCAGGGAATATCATCCAGACCATCAGCAGCTGGTAGAGT
ACAGCTTTCTCAATTGCTTTCCATGTTTTGGATACTTATATGCCCGTTAATAACAGGTAA
AATAGCCAGTACATCATTCCACATTTACCCATTGAATGTTGCATGTTTTCTTCTTTCC
CATATTCATACAGTCCAGATTTTCTTTTGGACTCATGACAGCACATTGGCTTTTCTTTCC
TTTCAGTTTCATGATTCTTAACCCCAAAGTGCTTTTGCCATGGGAACGGAAGGATAAATT

47666 GTGGTGATGGCTGGAATGTCAGGGGCAGGGGAGAGTACTGGCGTTAAGTTAAACCAACAG
ACATCCAGTTTAAACCACTGGTAGTTCTCAGTCTACATGTAGTTTATTTCTTCTGTTTATC
TGCCAATTTTATGTAGATCATCACATTGCCAAAAAAATCATTTTGAAGTGTATATAT
TTTTATGTGCATCATATTTATCTCCTAAATAAGTCTCTTCTTTCTTCTACTTTCTGATGCA
GACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTGGCTCGATGTTATAAC
[G, A]
TACGTGGGTCTACAAGTGCAAATGTAAGAAGGATACCAGGGTGTGGACTGACTTGTG
TGTGTGAGTAGCACTTGTCTCTCAGCTTTAAATTCTAGCAGGAAATACAGGATTACACAA
AGGCCATTGCTAGGGAAAATAAGGAATAAGATTATCAAAGAAGTATAATTGTCATAATTG
GTTATATTTGCTTTGATTTCCACAAACAATAAAATCACTTGCTCAGGTAATTTGTAATAA
CTAAGGACTCAGTAATACACTATAATCTTAAGAGTATTTAATCTCTTCACTGAAATCTC

52613 AGGTAATGGGACCTCTATCACCTGTTGCTAAGGTCTAATAATGTGGATTCACTCCTTAATT
ACTCTTCATACATACAGTCAATATGTAAGGAAATCCTGTTAGCTGTACCTTCAAATTATA
TTTAAGTGTGACCTTTTCTCACCACCTCCAGTGCTACCACCTGGCCCAAGCCATCTCTC
CCCTGGAATACAGGCCCAAAATCTTTTATCCAAATAACTTATGGAATATAGCATTTTTTA

FIGURE 3DD

GATTTTAGAAAATCAGTAAGGTACACATACTATATAGTACACACTGAAGTAGTGAACAT
[G, A]
CTGATTTTCTCTAGTGCCTTTACTGTGAACGTATCAATGTTAAGGAAAGGAAAATGATA
TTAACTCAGAGATGATGTCACAGCATATATTTACTAGCTTGCACAAATTTTTTAATGT
TAGCAAGATTTAAGGCAAAATTTTTACTTATATTTTAATTGGATCCTATGATGATTATTA
AAGAAAAAGATAGTTATCTTAGAAAGTATAAAATGAGCTCTTGGCAATGTGAGCAAAA
CAGAAAAATCAGATTTTTTTTAAAGTACGTGTACATTTGCTCATAGAGATTATAATTTAAAG

52645 GTCTAATAATGTGGATTTCACCTTAATTACTCTTCATACATACAGTCAATATGTAAGGAA
ATCCTGTTAGCTGTACCTTCAAATTATATTTAAGTGTGACCTTTTCTCACCAACTCCAGT
GCTACCACCCCTGGCCCAAGCCATCTCTCCCTGGAATACAGGCCCAAATCCTTTATCCA
AATAACTTATGGAATATAGCATTTTTTAGATTTTAGAAAATCAGTAAGGTACACATACTA
TATAGTACACACTGAAGTAGTGAAACATGCTGATTTTCCTCTAGTGCTTTTACTGTGAAC
[G, A]
TATCAATGTTAAGGAAAGGAAAATGATATTAACCTCAGAGATGATGTCTCACAGCATATAT
TTACTAGCTTGCACAAATTTTTTAATGTTAGCAAGATTTAAGGCAAATTTTTTACTTATA
TTTTAATTTGGATCCTATGATGATTATTAAGAAAAAAGTAGTTATCTCTAGAAAGTATAA
ATTGAGCTCTTGGCAAATGTGAGCAAAACAGAAATCAGATTTTTTTAAAGTTACGTGTA
CAATTTGTCTATAGAGTTATAATTTAAAGTATTTGTGCTCATAGCCAGTCTGTGTATCTGGG

59197 CTACTATTGATGTTGAATTCATTTATATAGTAATGATAACATTTCCTACTTAATTCATAA
AAAGACAGCCTATGCTGTTTCTTGTCTGAGTTTATATGTTTCTCATGCTTTTTATTAT
GGTTCATTACAATTTTAATGTTATTTTTAACTAACTAGATCCTTTTGAAACAAATTGGTT
TGCAAGTGTGAGCTGTTAGGTGCACAGAGAAAAATGAAAATAGAACTTGCGATTTTATT
CTAGGCTTGTACCAAATATTTAGAATACTGTGTTTTATTTAGGTGTTTATAGTCTCATT
[-, A]
GACAGTTGTGATTTTAAAATAGAGACCACATCATCTCAACTTCTTTACTGTGAAAAATAT
GACAATAGTCTTTTCAGAGATGAATCTGTCTAGATGGGAAATTTACATGATTGATCTGAT
GAGNN
AAAATAATCTGACAAGTAGTTTCCCCAGAAAATCTGATTTAGTAATGTACCAAAGGAT
TTAGAAATCTACATCATATAAATAAACATTCTATGTTATTTTAGTTTTCAGACCCATTTTAAT

63508 ACAGGCTCCCGTGCCAGGAAACCTTTACATCAACACTCGATTGCCATTTGATAGTCTT
TCATCTGGGAGGAAAAAAGACGGAGGGGAGCTTGAAAACTGTCATAATGTCCCTGG
AATATGGTACTTTTAAGAGTTGAGCCTATTCCATTTTGGAGATGATTTATATAAGTTACA
ACAAAAGAAGGGGACAAAAACATGATTGTTCTATGGAGTTTTTATAACTTTCTGTCACAA
GAAAGCACGCTTGTCTACAATTTTGTAATATTTCTAGTAAATAAAAGAGGCACCTCCCGT
[T, C]
TCAGAGCACCAAATAAGGAAAGTGTAATTGGATGTCATTGCTGTCAGTCAGCTGGGCTAT
AAAAGAGAGAGTGGGGTTGCCTCATCCCTGGGTATCCACAGTCAGCTGTGTCCCTAGAG
CTTCTTTCTTTTCATTGCTGCCAGCTGGGTATATTGCAAGTATGGATTATAAGAGGGGA
AGGGACTTCTAGCTTTTTACGTTTTGAAACAAAAAGGAAAAAAGCTAGAAGTAGTAAGCTA
AAAACAACCTTGTGCAACGTTCTGGGATTATTACTTAATTTTAAAGATTTTTGCTAAAAA

75153 AAGACATAGGTAGCAAAAAAATGGTCTAGGGCATCCCAACAAAATTATGAATGCATTTGG
GAAAATGTCATCATTTTTGCTGGTGCTTATGAAACTTGCTGTTGGGAGATACCTATGCCTG
TATGTAGGTATAGTTTATAAGAAACCATTTTAGAAAGAAAAGATAGATTCCATTAATACA
GTTTTTATCAGAGTACTTGGATTTTGTTTAATCTTACATATTTTTCTTAAAACTTTTC
TCAGTATTTTTATTGTTTAGAGAAATAAAACAAGATAATCATTTTAAATCATAGCACTTA
[C, G]
GTTTTCTCTTGTTTTATAAGGAGCAAGGATGCTCTATAGAAAATATAATGTAAGAATAAT
AAAAGTTTTTGGTTTTTACATAGGTAAAGCAACAGTGTGATTGGATTATGGTGTGTTGATT
CTATTCATCTTTTCAGACAAGAACGATGTTAACCACAAAGGAATGAAGAACATTTAAGA
TAGGCTGTGTGATTATGATCTTTTCAGATCTTTGGCTCCTAATCTGTTCCCTTTATATTC
TATCACACTCTTCTAACTTTGGTAATCCTTGACAAAAAGTGTGCATCTTTATAAACATACCT

76017 CGTCTTCAAAGGTGAAAAAGGCGTGCTCACACTGGGGAGATTGGATTAGATGATGTGAG
CTTGAAAAAAGGCCACTGCTCTGAAGAACGCTAACAACTCCAGAACTAACAAATGAAGTCC
TATGTTGCTCTATCCTCTTTTCCAATTCTCATCTTCTCTCCTCTTCTCCCTTTTATCAG
GCCTAGGAGAAGAGTGGGTCACTGGGTGAGAAGGAAGTCTATTTGGTGACCCAGGTTTTT

FIGURE 3EE

CTGGCCTGCTTTTGTGCAATCCCAATGAACAGTGATACCCTCCTTGAAATACAGGGGCAT
[C, G]
GCAGACACATCAAAGCCATCTGTGGGTGTTGCCCTCCATCCTGTGTCTCTTTCAGGAAGG
CATTACAGCATGCGTGAGCCATACCATCCTCCATCCTGATTACAAGGTGCTCCTTGAGCA
AATTATGAGAGTGAGTTACGGGAGCAGTTTTTAAAAGAAATCTTTCAGATGGCTATGAT
GTTATGTGTTTCGGTGTTGTACCATGAGTAGTATTGACTTCCCTTGAGATATGATGTACAA
TGTGCTTGTGAAATTGACTTACCCTCTTCACTTAAGTTAGTTCTGGCCTGACCTGAACTC

76033 AAAAGGCGTGGTCACACTGGGGAGATTGGATTAGATGATGTGAGCTTGAAAAAGGCCAC
TGCTCTGAAGAACGCTAACAACCTCCAGAATAACAATGAACCTCTATGTTGCTCTATCCT
CTTTTTCCAATTCTCATCTTCTCTCCTCTTCTCCCTTTTATCAGGCCTAGGAGAAGAGTG
GGTCAGTGGGTGAGAAGGAAGTCTATTTGGTGACCCAGGTTTTTCTGGCCTGCTTTTGTG
CAATCCCAATGAACAGTGATACCCTCCTTGAAATACAGGGGCATCGCAGACACATCAAAG
[C, T]
CATCTGTGGGTGTTGCCCTCCATCCTGTGTCTCTTTCAGGAAGGCATTACAGCATGCGTGA
GCCATACCATCCTCCATCCTGATTACAAGGTGCTCCTTGAGCAAATTATGAGAGTGAGT
TACGGGAGCAGTTTTTAAAAGAAATCTTTCAGATGGCTATGATGTTATGTGTTTCGGTGT
TGTACCATGAGTAGTATTGACTTCCCTTGAGATATGATGTACAATGTGCTTGTGAAATTG
ACTTACCCTCTTCACTTAAGTTAGTTCTGGCCTGACCTGAACTCTGACTTTTACTGCCAT

77194 AACGTTTTAAATTGATATAAGTTTAGGCAGTTGTAGTTCATAACTTATGTTGCTCATGTT
GTGCTGTGTAGGATGGGATAGGAAGCAAGTCCCATGCTTAGAGGCATGGGATGTGTTGG
AACGGGATTACACACACTGGAGGAGCAGGGCAAGTTGGAATTCTAAGATCCATGAACCC
CCAAGTGTATTTCTCCCTGCATATTTTACCAATATATTAACCAATGTAACCTTTTAA
AAGGCATCATTCCTGAGGTTTGTCTTAATTTCTGATTAAGTAATCAGAATATTTTCTGCT
[A, G]
TTTTTGCCAGGAATCACAAAGATGATTAAAGGGTTGGAAAAAAGATCTATGATGGAAAA
TTAAAGGAAGTGGGATTATTGAGCCTGGAGAAGAGAAGACTGAGGGGCAAACCATTTGATG
GTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTGTTCTCCATATGC
ACTAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGC
AGGGGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTC

77508 TCACAAAGATGATTAAAGGGTTGGAAAAAAGATCTATGATGGAAAATTAAAGGAAGTGG
GATTATTGAGCCTGGAGAAGAGAAGACTGAGGGGCAAACCATTTGATGGTTTTCAAGTATA
TGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTGTTCTCCATATGCACTAAGAATAGAA
CAAGAGGAAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGGGGCCATTGTT
AGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTTTCTAAAAAA
[T, A]
TAGATAAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACA
AATTATAGAATTTCCCAAAAGATGTTTGTATCTACTAGTAGTATGCAGTGAAAATCTTT
AGAACTAAATAATTTGGACAAGGCTTAATTTAGGCATTTCCTCTTGACCTCCTAATGGA
GAGGGATTGAAAGGGGAAGAGCCCAAAATGCTGAGCTCACTGAAATATCTCTCCCTTA
TGGCAATCCTAGCAGTATTAAAGAAAAAAGGAACTATTTATTCCAATGAGAGTATGAT

77557 AAAGGAAGTGGGATTATTGAGCCTGGAGAAGAGAAGACTGAGGGGCAAACCATTTGATGGT
TTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTGTTCTCCATATGCAC
TAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAG
GGGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTC
TTTCTAAAAAATTAGATAAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAA
[G, A]
GAAAAGTAACAAATTATAGAATTTCCCAAAAGATGTTTGTATCTACTAGTAGTATGCAG
TGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGGCATTTCCTCTTGAC
CTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCAAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAACTATTTATTCCAAT
GAGAGTATGATGGACAGATATTTTAGTATCTCAGTAATGTCTAGTGTGGCGGTGTTTT

FIGURE 3FF